NASA TECHNICAL NOTE

NASA TN D-6571

APPLICATION OF MONTE CARLO TECHNIQUES TO OPTIMIZATION OF HIGH-ENERGY BEAM TRANSPORT IN A STOCHASTIC ENVIRONMENT

by Russell V. Parrish, James E. Dieudonne, and Tassos A. Filippas

Langley Research Center
Hampton, Va. 23365

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • DECEMBER 1971

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TN D-6571		

4. Title and Subtitle	5. Report Date
APPLICATION OF MONTE CARLO TECHNIQUES TO OPTIMIZATION OF HIGH-ENERGY BEAM TRANSPORT IN A STOCHASTIC ENVIRONMENT	December 1971
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Russell V. Parrish, James E. Dieudonne, and Tassos A. Filippas	L-7930
	10. Work Unit No. 136-62-01-03
9. Performing Organization Name and Address	11. Contract or Grant No.
NASA Langley Research Center Hampton, Va. 23365	
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Note
National Aeronautics and Space Administration Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

An algorithm employing a modified sequential random perturbation, or creeping random search, has been applied to the problem of optimizing the parameters of a high-energy beam transport system. The stochastic solution of the mathematical model for first-order magnetic-field expansion allows the inclusion of state-variable constraints, and the inclusion of parameter constraints allowed by the method of algorithm application eliminates the possibility of infeasible solutions. The mathematical model and the algorithm were programed for a real-time simulation facility; thus, two important features are provided to the beam designer: (1) a strong degree of man-machine communication (even to the extent of bypassing the algorithm and applying analog-matching techniques), and (2) extensive graphics for displaying information concerning both algorithm operation and transport-system behavior. Chromatic aberration was also included in the mathematical model and in the optimization process.

Results presented show this method as yielding better solutions (in terms of resolutions) to the particular problem than those of a standard analog program as well as demonstrating flexibility, in terms of elements, constraints, and chromatic aberration, allowed by user interaction with both the algorithm and the stochastic model. Also presented are examples of slit usage and a limited comparison of predicted results and actual results obtained with the 600 MeV cyclotron at the NASA Space Radiation Effects Laboratory.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Beam transport optimization Stochastic optimization Beam transport Optimization with constraints	Unclassified – Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	46	$3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

APPLICATION OF MONTE CARLO TECHNIQUES TO OPTIMIZATION OF HIGH-ENERGY BEAM TRANSPORT IN A STOCHASTIC ENVIRONMENT

By Russell V. Parrish, James E. Dieudonne,
and Tassos A. Filippas
Langley Research Center

SUMMARY

An algorithm employing a modified sequential random perturbation, or creeping random search, has been applied to the problem of optimizing the parameters of a high-energy beam transport system. The stochastic solution of the mathematical model for first-order magnetic-field expansion allows the inclusion of state-variable constraints, and the inclusion of parameter constraints allowed by the method of algorithm application eliminates the possibility of infeasible solutions. The mathematical model and the algorithm were programed for a real-time simulation facility; thus, two important features are provided to the beam designer: (1) a strong degree of man-machine communication (even to the extent of bypassing the algorithm and applying analog-matching techniques), and (2) extensive graphics for displaying information concerning both algorithm operation and transport-system behavior. Chromatic aberration was also included in the mathematical model and in the optimization process.

Results presented show this method as yielding better solutions (in terms of resolutions) to the particular problem than those of a standard analog program and well as demonstrating flexibility, in terms of elements, constraints, and chromatic aberration, allowed by user interaction with both the algorithm and the stochastic model. Also presented are examples of slit usage and a limited comparison of predicted results and actual results obtained with the 600 MeV cyclotron at the NASA Space Radiation Effects Laboratory.

INTRODUCTION

With the advent of large-accelerator projects over the past 15 years, a demand has arisen for optimization of not only the design of the accelerator but also the design of the beam-handling systems that must necessarily be coupled with the accelerator. As pointed out in reference 1, the high cost of producing the strong magnetic fields over large volumes necessary for high-energy beam optics makes optimization a necessity.

Past efforts in the area of beam-transport optimization have generally fallen into two classes of techniques: "analog matching" with a high-speed repetitive analog computer and iterative deterministic methods for parameter optimization implemented on a general-purpose digital computer. References 1 to 3 describe the general methods used in the technique of analog matching: Generally in designing transport systems the results of parameter variations in the system are continuously displayed on an oscilloscope as the designer seeks by trial and error to determine the focusing strengths and separation distances necessary to give the desired results. Aside from the problems of analog resolution and repeatability, which are usually present even with advanced equipment, the major disadvantage with this technique is that it relies heavily on the experience of the analyst, especially as the number of parameters increases. It is extremely time consuming and is usually based on visual determination of improvement; thus, the problem of visual resolution arises.

However, as references 1 to 3 strongly emphasize, the man-machine communication available with the analog computer is often felt to be the overriding consideration when choosing between the two classes of techniques available to solve the problem of beam-transport design. This man-machine communication is generally sacrificed when the analyst chooses to use a digital-computer optimization program. This sacrifice is not unwarranted, however, according to references 4 and 5, which emphasize the advantages of optimization algorithms over analog methods. These advantages include a performance measure for evaluating a given set of parameters, an algorithm that automatically seeks an extremum solution of the performance measure (which may or may not be feasible), and digital resolution.

Thus a survey of the literature available in the field reveals strong advocates for each class of techniques. The present paper will describe a method that evolved in the process of solving a particular problem in beam-transport design. This method combines some of the best features of both classes of techniques, that is, a digital algorithm with its ensuing resolution, and man-machine communication (complete communication even to the extent of bypassing the algorithm and applying analog-matching techniques). In addition, the method allows for the inclusion of state-variable constraints and parameter constraints in the mathematical model. Other features of the program include stochastic representation of initial conditions and chromatic aberration and the introduction of the aberration during the optimization process. These features make the simulation of the beam transport system more realistic than in most previous work (e.g., ref. 1).

The method will be presented in the context of solving a specific problem in beam-transport design, namely, that of transporting a beam of charged particles from one given area to another through the use of four quadrupole magnets (other configurations of ele-

ments are allowable). In this problem five parameters are available for the optimization process. The maximum number of parameters that can be varied at one time is determined by the choice of the optimization algorithm and the amount of computing time necessary to evaluate the performance measure. However, the method may be used to optimize several elements at a time along the transport system, with the resulting beam used as input to the next set of elements in large systems.

SYMBOLS

a_i, b_i	constraint constants, $meter^{-1}$
C	value of cost function with present parameter values, dimensionless
C_p	value of cost function at last success, dimensionless
C_x	cost function corresponding to horizontal axis, dimensionless
C_z	cost function corresponding to vertical axis, dimensionless
D2	separation distance between second and third quadrupoles, meters
e	electron charge, coulombs
F_{cal}	calculated test statistic, dimensionless
$F_{.95}(n,n_p)$	F value for 95-percent confidence level with (n,n_p) degrees of freedom, dimensionless
g_i	ith magnetic-field gradient, $webers/meter^2$
H_A	alternate hypothesis for a statistical test, dimensionless
H_0	null hypothesis for a statistical test, dimensionless
K_i	magnetic focusing strength of ith quadrupole, $meter^{-1}$
l_i	effective field width, meters
m	number of rays which passed through the system, dimensionless

n	number of rays transmitted with present parameter values, dimensionless
n_p	number of rays transmitted at last success, dimensionless
N	percentage of rays which did not pass through the system, $\frac{n - m}{n} \times 100$, dimensionless
$P_{f,x}$	horizontal focal point, meters
$P_{f,z}$	vertical focal point, meters
$(\Delta P_f)_{min}$	minimum value allowed for distance between horizontal and vertical focal points of the first doublet, meters
$(P_{f,x})_{min}$	minimum value allowed for horizontal focal point of the first doublet, meters
$(P_{f,x})_{max}$	maximum value allowed for horizontal focal point of the first doublet, meters
$(P_{f,z})_{min}$	minimum value allowed for vertical focal point of the first doublet, meters
$(P_{f,z})_{max}$	maximum value allowed for vertical focal point of the first doublet, meters
p_o	average momentum, kilogram-meters/second
Δp	momentum perturbation, kilogram-meters/second
r	thickness of slit, meters
s	displacement along focal axis of transport system, meters
Δs	$s_{out} - s_n$, meters
$s_{e,x}$	position along focal axis of extremum in x, meters
$s_{e,z}$	position along focal axis of extremum in z, meters
s_f	position along focal axis at which cost function is evaluated, meters

$s_{h,B}$ beginning position of horizontal slit, meters

$s_{h,E}$ ending position of horizontal slit, meters

s_n position of entrance to field (magnetic or drift) along focal axis, meters

s_{out} position of exit from field (magnetic or drift) along focal axis, meters

$s_{v,B}$ beginning position of vertical slit, meters

$s_{v,E}$ ending position of vertical slit, meters

t_{cal} calculated test statistic, dimensionless

t_o t value associated with type I error, dimensionless

t_1 t value associated with type II error, dimensionless

$t_{.05}\left(n + n_p - 2\right)$ Student's "t" for 95-percent confidence level with $\left(n + n_p - 2\right)$ degrees of freedom, dimensionless

W_{dif} weight corresponding to difference between cost function for vertical plane and cost function for horizontal plane, dimensionless

W_N weight corresponding to percentage of beam which passed through the system, dimensionless

$x(s)$ displacement along horizontal axis of transport system, meters

x_d desired value of horizontal displacement at end point, meters

$\overline{x'}\left(s_f\right)$ mean value of the horizontal slope at the end point of all rays of the beam which passed through the system, dimensionless

x_{sl} displacement of beam in horizontal plane parallel to focal axis, used to determine slit opening, meters

$z(s)$ displacement along vertical axis of transport system, meters

z_d desired value of vertical displacement at end point, meters

$\overline{z'}(s_f)$ mean value of the vertical slope at the end point of all rays of the beam which passed through the system, dimensionless

α_i aperture radius, meters

δ detectable difference desired, dimensionless

σ^2 variance of cost function for present parameter values, dimensionless

$\hat{\sigma}_p^2$ estimate of cost-function variance at last success, dimensionless

σ_p^2 variance of cost function at last success, dimensionless

$\hat{\sigma}^2$ estimate of cost-function variance for present parameter values, dimensionless

ω_x slit width along x-axis, meters

ω_z slit width along z-axis, meters

Subscript:

d desired value

Primes indicate derivatives with respect to the focal axis (s-axis).

MATHEMATICAL MODEL

The basic beam-transport problem discussed in this paper is the design of a high-energy beam transport system consisting of four quadrupoles, with variable magnetic focusing strengths and separation distances, and slits with variable width, thickness, and separation distance. This system will transform a charged-particle beam of known initial conditions into a beam meeting some desired phase-plane configuration defined by a mathematical formula known as the cost function. In the design of this particular system, illustrated in figure 1, five system parameters were to be determined. These were the four magnetic focusing strengths $(K_1, K_2, K_3, \text{and } K_4)$ and the separation distance D_2

between the doublets (pairs of quadrupoles). All other system parameters are considered fixed. In general, the method used to determine the parameters imposes no restrictions as to number or types of the parameters.

For applications of interest here, the differential equations of motion for traversing a quadrupole field (ref. 1, p. 9)

$$\left.\begin{aligned} x'' &= -K_i\left(1 + z'^2 + x'^2\right)^{1/2}\left[\left(1 + x'^2\right)x - x'z'z\right] \\ z'' &= K_i\left(1 + z'^2 + x'^2\right)^{1/2}\left[\left(1 + z'^2\right)z - x'z'x\right] \end{aligned}\right\} \tag{1}$$

can be reduced to the uncoupled, linear equations (refs. 1 and 6)

$$\left.\begin{aligned} x'' + K_i x &= 0 \\ z'' - K_i z &= 0 \end{aligned}\right\} \tag{2}$$

where K_i, a constant (i = 1,2,3,4), is the focusing strength of the field. It represents a rectangular, first-order, field model $K_i = eg_i/p_o$ (where the ith magnetic potential is equal to $-g_i \cdot xz$) in which focusing occurs in the xs-plane and defocusing in the zs-plane. The algebraic solutions of equations (2) are shown in the following equations:

$$x\left(s_{out}\right) = x\left(s_n\right)\cos\left(\sqrt{K_i}\,\Delta s\right) + \frac{x'\left(s_n\right)\sin\left(\sqrt{K_i}\,\Delta s\right)}{\sqrt{K_i}} \tag{3a}$$

$$x'\left(s_{out}\right) = -x\left(s_n\right)\sqrt{K_i}\,\sin\left(\sqrt{K_i}\,\Delta s\right) + x'\left(s_n\right)\cos\left(\sqrt{K_i}\,\Delta s\right) \tag{3b}$$

$$z\left(s_{out}\right) = z\left(s_n\right)\cosh\left(\sqrt{K_i}\,\Delta s\right) + \frac{z'\left(s_n\right)\sinh\left(\sqrt{K_i}\,\Delta s\right)}{\sqrt{K_i}} \tag{3c}$$

$$z'\left(s_{out}\right) = z\left(s_n\right)\sqrt{K_i}\,\sinh\left(\sqrt{K_i}\,\Delta s\right) + z'\left(s_n\right)\cosh\left(\sqrt{K_i}\,\Delta s\right) \tag{3d}$$

$$\Delta s = s_{out} - s_n$$

The inverse case, namely, focusing in the zs-plane and defocusing in the xs-plane, may be implemented by using equations (2) as shown, but with the sign of K_i changed.

The drift-space translational equations

$$\left.\begin{aligned} x'' &= 0 \\ z'' &= 0 \end{aligned}\right\} \tag{4}$$

also have algebraic solutions:

$$x(s_{out}) = x(s_n) + x'(s_n)\Delta s \tag{5a}$$

$$x'(s_{out}) = x'(s_n) \tag{5b}$$

$$z(s_{out}) = z(s_n) + z'(s_n)\Delta s \tag{5c}$$

$$z'(s_{out}) = z'(s_n) \tag{5d}$$

Throughout the entire translation along the s-axis, the position variables are subject to the physical constraints imposed by the aperture size of each successive quadrupole:

$$x^2(s) + z^2(s) \leqq \alpha_i^2 \tag{6}$$

where α_i are the aperture radii for the four quadrupoles (i = 1,2,3,4). This constraint was imposed at the entrance and exit of each quadrupole as well as within the quadrupole. The locations of the extremum values of x and z $\left(s_{e,x} \text{ and } s_{e,z}\right)$ within the focusing quadrupole are determined by

$$\left.\begin{aligned} s^*_{e,x} &= \frac{1}{\sqrt{K_i}} \tan^{-1} \frac{x'(s_n)}{\sqrt{K_i}\,x(s_n)} \\ s^*_{e,z} &= \frac{1}{\sqrt{K_i}} \tanh^{-1} \frac{z'(s_n)}{\sqrt{K_i}\,z(s_n)} \end{aligned}\right\} \tag{7}$$

where

$$s_{e,x} = \begin{cases} 0 & \left(s^*_{e,x} < 0\right) \\ s^*_{e,x} & \left(0 \leqq s^*_{e,x} \leqq \Delta s\right) \\ \Delta s & \left(s^*_{e,x} > \Delta s\right) \end{cases} \qquad s_{e,z} = \begin{cases} 0 & \left(s^*_{e,z} < 0\right) \\ s^*_{e,z} & \left(0 \leqq s^*_{e,z} \leqq \Delta s\right) \\ \Delta s & \left(s^*_{e,z} > \Delta s\right) \end{cases}$$

Again, the equations presented are for focusing in the xs-plane, and a change in sign of K_i is necessary for focusing in the zs-plane.

In addition to the position constraints of equation (6), two types of parameter constraints may be present in the problem. The first type

$$\left.\begin{aligned} -b_i \leqq K_i \leqq b_i \qquad & (i = 1,2,3,4) \\ a_1 \leqq D_2 \leqq a_2 \qquad & \end{aligned}\right\} \tag{8}$$

is imposed by the physical properties of the quadrupoles to be used in the actual experiment, limiting the values of K_i to feasible ranges and the separation distance between the doublets to values meeting the requirements of the experimental laboratory dimensions.

The second type of parameter constraint is introduced by the desire to use slits to improve the beam quality. Slits are usually to be introduced at some focal point in either the xs- and zs-planes. It is desirable that the focal points fall within the drift space between the elements and still allow enough room for slit placement. It is also desirable that the focal points be separated from one another by sufficient distance for placement of slits of different thickness at each focal point. These constraints are represented by

$$\left(P_{f,x}\right)_{min} \leqq P_{f,x} \leqq \left(P_{f,x}\right)_{max} \tag{9a}$$

$$\left(P_{f,z}\right)_{min} \leqq P_{f,z} \leqq \left(P_{f,z}\right)_{max} \tag{9b}$$

$$\left|P_{f,x} - P_{f,z}\right| \geqq \left(\Delta P_f\right)_{min} \tag{9c}$$

and the focal points are determined by solving equations (3) with initial conditions $x' = z' = 0$, and $x = z = \text{Constant}$ through the desired elements. The focal points are then determined by the points of intersection with the s-axis.

Slits are implemented as position constraints:

$$\left.\begin{aligned} |x(s)| &< \frac{\omega_x}{2} \qquad \left(s_{h,B} \leqq s \leqq s_{h,E}\right) \\ |z(s)| &< \frac{\omega_z}{2} \qquad \left(s_{v,B} \leqq s \leqq s_{v,E}\right) \end{aligned}\right\} \tag{10}$$

assuming no penetration (complete absorption), and can be introduced separately, jointly, or not at all. With the above constraints imposed on the system, the problem is to find a set of parameters which when placed in the system will yield a "best fit" to some predetermined phase-space curve defined in the cost function. The algorithm chosen to perform this optimization is explained in the section of this paper entitled "Monte Carlo Method."

One primary aberration, chromatic aberration, was taken into account in the mathematical model. This aberration arises from the different values of momentum possessed by the individual rays. In the equations presented previously, the momenta of the particles are considered constant and equal throughout the transport system, that is, $K_i = eg_i/p_o$. Usually an optimal case is obtained without the presence of chromatic aberration, and then the effect of the aberration is examined in the following manner: The varying momentum is modeled as $K_i = \dfrac{eg_i}{p_o + \Delta p}$, where Δp is a positive or negative momentum perturbation. In the present paper, each ray of the beam is assigned a momentum perturbation based on Gaussian noise. The perturbation is then applied to each focusing strength. In this manner, the momentum of an individual particle of the beam is constant throughout the focusing system, while individual momenta within the beam vary. Optimization then takes place in the presence of chromatic aberration.

ANALOG SIMULATION

During the preliminary investigation of the problem, the equations of motion (eqs. (2)) and the constraint equations (eqs. (8)), were programed on both the GPS 10000 repetitive analog computer and the EAI 680 analog-hybrid computer. The function K_i, which represents the system parameters, was generated on both machines by the use of

electronic comparators and available logic components. The variation of system parameters was achieved by tandem potentiometers. Repetitive solutions of the differential equations at a rate of 10 runs per second allowed oscilloscope displays of the trajectories z(s) and x(s) (beam traces) and phase-plane plots in both planes at any point in the system.

A visual-resolution problem was encountered in the use of this method in that the operator found many promising solutions, or sets of system parameters, from which by visual means a "best case" could not be determined. To solve this problem, a digital program was written for the solution of the differential equations with the predetermined sets of system parameters. The digital program, which also served as an independent check, solved the resolution problem, but uncovered the fact that the analog-determined system parameters did not fulfill the desired phase-plane end conditions. This fact will be demonstrated in the "Results" section of this paper.

REAL-TIME SIMULATION SYSTEM

The inability of the analog program in meeting required phase-plane end conditions led to the programing of the problem on the CDC 6600 Real Time Simulation (RTS) Subsystem at the Langley Research Center. This system, which is described in reference 7, provides both sufficient accuracy for the problem and man-machine communication. In addition to the main computers, the RTS system contains program control stations, an analog-to-digital and discrete input system, a digital-to-analog and discrete output system, a CDC 250 Series CRT display system, and peripheral units (time-history recorders, x-y plotters, card readers, etc). The RTS software provides many "analog-like" features such as mode controls, subroutines for displaying and changing parameters, and function-sense switches.

The application of the RTS system to the beam-transport problem allows the use of an optimization algorithm to determine an optimal set of system parameters while the operator maintains man-machine communication.

This communication link to the machine allows quick, effective use of the designer's knowledge of the system. It enables him to examine different areas of the parameter space and determine the merits of particular parameter sets as to likely starting points for the optimization algorithm, thereby reducing the time required for convergence. The designer also has many options available through subroutines, function-sense switches, and the RTS software from which he may change the model of the transport system, position-variable and parameter constraints, the form of the desired output of the system, and the form of the probability distribution that dictates the initial conditions of the beam prior to magnetic-field entrance.

Monte Carlo Method

The individual rays of the incident beam in the model are chosen to represent the physical beam for which the transport system is being designed. This is done by choosing a set of random initial conditions that satisfies the physical requirements. Equations (3) are then evaluated to provide the values of the position variables at the entrance and exit of each magnetic field, and also the extremum values within the magnetic fields are evaluated (eqs. (7)). Any violations of the position-variable constraints are recorded (solution of the equations of motion for an individual ray is terminated upon the occurrence of a violation) and may be included in a weighted performance measure with the deviation of the transmitted particles from the desired end conditions. This process is then repeated and the deviations are accumulated to form the mean performance measure, or mean cost function, over the desired sample size. The parameters of the system (four focusing strengths and one separation distance) are then perturbed by the optimization algorithm and the entire sequence is iterated until convergence is achieved.

Monte Carlo Algorithm

The optimization algorithm used in conjunction with the mathematical model was a variation of the sequential random perturbation, or creeping random search, suggested by Mitchell in reference 8. This algorithm was selected over deterministic methods that are gradient dependent, such as steepest descent, conjugate gradient, Newton-Raphson, and quasilinearization (see ref. 9), because of the presence of constraints and a stochastic model. The algorithm is described in detail in reference 10.

A stochastic representation of the initial beam (beam prior to entry into the transport system) was advantageous in making the simulation of the system more realistic. By considering the beam as a conglomerate of rays, each having individual characteristics of displacement, slope, and momentum, position-variable constraints may be implemented, and the percentage of the beam striking the elements of the transport system may be calculated. The stochastic model also allows for realistic representation of chromatic aberration since each ray of the beam instead of the beam as a whole is assigned a momentum perturbation.

The basic algorithm, devoid of any strategies, is best illustrated with the two-parameter contour plot shown in figure 2(a). In this simple problem, constant values of the cost function are represented as concentric circles in the two-parameter plane. From an arbitrary starting point, a random step is taken and the cost function is evaluated. If the step has improved the cost function (a success), the present parameter values are used as a new starting point and a new step is taken. If the cost function was not improved (a failure), the original parameter values are retained, and a new step is taken. Thus a "walk" over the contour is generated.

For explanation of the basic algorithm in more detail, attention is directed to figure 2(b). As illustrated in the figure, there are nine possible directions that can be taken from a given starting point in the two-parameter plane. These nine directions are obtained by plus, minus, and zero perturbations of each parameter. (Note that one direction is no movement at all.) The eight directions that involve movement are stored in an array, which the algorithm will sample randomly without replacement. The array is sampled until a success is obtained or all directions are exhausted, at which point the array is reinitialized. If failures should occur in all possible directions from a given point, a reduction in step size would be necessary in order to continue iteration. The algorithm is stopped once the step size becomes smaller than the desired detectable difference (see eq. (13)) between different values of the cost function.

In conjunction with the basic algorithm, two features are available to speed convergence. The first feature is a strategy implemented at the option of the designer by the algorithm and simply requires that after each success the next try be made in the same direction. The second feature involves interaction of the beam designer with the algorithm through the CRT displays and will be described later.

Parameter Constraints

The parameter constraints of equations (8) are implemented by simply not allowing the algorithm to perturb parameters across boundaries. Thus these constraints affect the problem only in that the parameters must always be within their feasible ranges. The focal-point constraints, equations (9), are implemented by perturbing parameters, calculating the focal points, and sensing for violations. Should a violation occur, the perturbation direction is deemed a failure, the equations of motion are not evaluated, and a new direction is chosen by the algorithm. It should be obvious that the addition of the focal-point constraints greatly speeds convergence, since the evaluation of the equations of motion is bypassed for all parameter sets which do not satisfy the constraints.

Statistical Test for Success

Assuming equal variances, a pooled t test for equality of means, as described in reference 11, is used to determine the success or failure of a perturbation direction. The null hypothesis was chosen to assure control over the probability of accepting a failure as a success (a 5-percent chance of occurrence was selected). The test is shown as follows:

H_0: $C_p \leqq C$ (Failure) $\qquad$ H_A: $C_p > C$ (Success)

$$t_{cal} = \frac{C_p - C}{\sqrt{\left(\frac{1}{n_p} + \frac{1}{n}\right)\frac{(n_p - 1)\hat{\sigma}_p^2 + (n - 1)\hat{\sigma}^2}{n_p + n - 2}}} \tag{11}$$

Reject H_0 if $t_{cal} > t_{.05}(118) = 1.66$

where $n_p + n - 2 = 118$ degrees of freedom. Here, C_p is the value of the mean cost at the last success; n_p is the sample size (60 rays/beam) at the last success; $\hat{\sigma}_p^2$ is the variance at the last success; and C, n, and $\hat{\sigma}^2$ are the mean cost, sample size, and variance, respectively, for the present parameter values being tested. It should be noted that an F test was used periodically to validate the assumption of equality of variances, as shown in the following equations:

$$H_0: \ \sigma_p^2 = \sigma^2 \qquad H_A: \ \sigma_p^2 \neq \sigma^2$$

$$F_{cal} = \frac{A}{B} \quad \text{where} \begin{cases} A = \max\left(\hat{\sigma}_p^2, \hat{\sigma}^2\right) \\ B = \min\left(\hat{\sigma}_p^2, \hat{\sigma}^2\right) \end{cases} \tag{12}$$

$$\text{Reject } H_0 \text{ if } F_{cal} > F_{.95}(60,60) = 1.67$$

The null hypothesis was accepted in all cases, except for a few in which the step size of the perturbations was large.

The sample size (60 rays/beam) was determined by using the iterative method described in reference 12, pages 154-155,

$$n \geq \frac{2\left(t_o + t_1\right)^2 \hat{\sigma}^2}{\delta^2} \tag{13}$$

where

t_o t value associated with type I error; $t_o = t_{.05}(118) = 1.66$

t_1 t value associated with type II error (95-percent power for the chosen detectable difference); $t_1 = t_{.10}(118) = 1.29$

$\hat{\sigma}^2$ estimate of variance; $\hat{\sigma}^2 \approx 0.01$

δ detectable difference desired; $\delta = 0.055$

GRAPHIC DISPLAYS AND INTERACTIVE CAPABILITIES

Extensive use of the CRT displays is an integral part of program operation. In addition to the algorithm displays, the following problem-related displays are available to the analyst:

(1) Phase plots for both the x and z variables at the entrance to the transport system and at the exit of each quadrupole

(2) An xz-plane plot at the same locations

(3) Beam traces, both horizontal and vertical views, including the location and size of each quadrupole, slit placement in the system, and the horizontal and vertical focal points

Thus the analyst has complete information of the transport system at important points along the system. If the designer desires, based on his knowledge of transport systems, algorithm operation may be bypassed, and with use of the displays, analog matching may be carried out. This "hands-on" capability allows the designer to look at areas in the parameter contour which he may deem interesting.

The algorithm displays are available to possibly speed convergence to an extremum of the problem in the following manner: The designer may, during algorithm operation, display on the CRT a history of the cost-function behavior versus the number of tries (both successes and failures) and also histories of the value of each parameter versus the number of successes. Based on this information, the designer may be able to determine the sensitivities of certain parameters in the current region of operation of the cost domain. If the parameter appears sensitive, the gain, or perturbation size of that parameter, may be increased. On the other hand, if a certain value is required of a parameter in order to obtain a success or if the parameter appears insensitive (varies randomly about some mean), it may be made a constant. In this situation, the number of possible directions is greatly reduced, as evidenced by table 1, and the array of directions is regenerated for the reduced number of parameters. Therefore, increasing the gain of sensitive parameters and eliminating insensitive parameters should then speed convergence to an optimal set of parameters.

RESULTS

The major fields of interest in beam-transport design are transporting and transforming either a parallel source or a point source, at the exit of the accelerator, into a parallel distribution or a point distribution at a desired location. Thus, four cases arise: parallel to parallel, parallel to point, point to parallel, and point to point. All four are presented to illustrate the versatility of the method.

Figure 3 shows a typical sampling from the initial distribution used in all parallel-source examples except the parallel-to-point case. Although 60 points were used in each algorithm iteration, only 30 are shown in order to provide clarity in the plotting. The distribution used to model the beam-entrance conditions was fixed along the envelope of

the phase-plane area, since interior points in the phase-plane area remain in the interior throughout the transport system.

Four studies were made under the parallel-to-parallel case, all of which used the equations

$$C_X = \frac{\sum_{i=1}^{m}\left[\frac{x(s_f)_i}{x_d(s_f)}\right]^2 + \left[\frac{x'(s_f)_i}{x'_d(s_f)}\right]^2}{m} \tag{14a}$$

$$C_Z = \frac{\sum_{i=1}^{m}\left[\frac{z(s_f)_i}{z_d(s_f)}\right]^2 + \left[\frac{z'(s_f)_i}{z'_d(s_f)}\right]^2}{m} \tag{14b}$$

$$C = C_X + C_Z + \left|C_X - C_Z\right| W_{dif} + W_N N \tag{14c}$$

as the cost function defining the desired end conditions. (The desired end conditions are represented as an ellipse in the appropriate phase-plane figures.) The four studies are (1) best analog results, (2) best digital results, (3) addition of chromatic aberration to the digital problem, and (4) addition of slits to the digital problem.

Table 2 together with figures 4 to 7 shows a comparison of typical results at the end of the last quadrupole obtained from the analog program and the digital program using the analog-obtained parameter values as a starting point for the digital program. As exemplified in table 2, the digital program in all cases compared made changes in the parameters which lowered the value of the cost function. Comparison of figures 5 and 7 reveals the visual-resolution problem mentioned previously. Figures 8 and 9 are results obtained by adding 5 percent chromatic aberration to the digital system of figures 6 and 7. Table 2 also lists pertinent information for this case. During this study it was noted that in most cases the introduction of chromatic aberration did not affect the convergence of the algorithm.

Table 3 and figures 10 to 13 illustrate an example of the introduction of a slit in one plane to achieve some desired end conditions which cannot be met by the system without slits. In the example the requirement to eliminate all slopes greater than a certain value in the vertical plane, as shown by the points outside the ellipse in the lower portion of figure 10, has been placed on the system. Assuming it has been determined that the system, subject to parameter constraints, cannot meet the requirement, the problem is solved in

the following manner: First, the additional constraint that the focal point of the first doublet in the vertical plane lie within a certain area between the two doublets is introduced into the system (eq. (9b)). The algorithm is then used to determine a set of system parameters satisfying all constraints which yield end conditions best fitting the imposed requirement. At this time a slit is placed at the vertical focal point of the first doublet, and the slit opening is varied until the imposed requirement is met and the desired end conditions are achieved (fig. 12).

The last parallel-to-parallel case is a comparison of computer-generated results with those obtained from an actual experiment conducted at the NASA Space Radiation Effects Laboratory. Figure 14(a) reflects the computer-generated results while figure 14(b) depicts the results obtained from the actual experiment. (Fig. 14(b) shows the proton beam passing through a spark chamber after it has exited the beam transport system.)

In the point-to-parallel study, equations (14) are still used as the cost function since the desired end conditions are the same as in the parallel-to-parallel case. Therefore, the only change required to enable the program to go from the parallel-to-parallel case to the point-to-parallel case is a change in initial conditions. The algorithm will then determine a new set of parameters which will best meet desired end conditions. Figure 15 shows a typical sampling from the initial distribution used to simulate a point source. Figure 16 reflects the end conditions obtained, and table 4 lists the set of system parameters obtained by the program in addition to values of other variables for the point-to-parallel case. Figure 17 gives a good visual illustration of the point-to-parallel case.

The parallel-to-point and point-to-point cases are obtained by using the appropriate initial conditions together with a cost function defining the desired end conditions as a point. The following equations represent the cost function used in these studies:

$$C_X = \frac{\sum_{i=1}^{m} \left[x\left(s_f\right)_i \right]^2}{m} \tag{15a}$$

$$C_Z = \frac{\sum_{i=1}^{m} \left[z\left(s_f\right)_i \right]^2}{m} \tag{15b}$$

$$C = C_X + C_Z + \left| C_X - C_Z \right| \left(W_{dif} \right) + W_N N \tag{15c}$$

Figure 18 is a typical sampling of the initial distribution used for the parallel-to-point case. Figure 19 and table 5 show the results obtained from the program. It should be noted that the distance D_4 between the end point and the beginning of the first magnet is an additional parameter to be determined in both the parallel-to-point and point-to-point cases.

The final case, point to point, is shown in figure 20 with its pertinent data given in table 6. The initial conditions used in this case are the same as for the point-to-parallel case shown in figure 15.

CONCLUDING REMARKS

A modified creeping-random-search algorithm has been applied successfully to solving a particular beam-transport design problem. The method of application that evolved in the process of solving this problem possesses the advantages of algorithm operation with analyst interaction, or if desired, analog matching with digital accuracy, and feasible solutions in the presence of both position-variable and parameter constraints. In addition, realistic modeling of chromatic aberration has been included in the mathematical model as affecting individual members of the beam, rather than the beam as an entity.

The ability of the method to handle all major cases has been demonstrated using a particular mathematical model, although flexibility in the model is possible. Results presented have shown that this method not only yields better solutions than those of a standard analog program, but also yields solutions for systems, subject to constraints, for which prior transport optimization programs give infeasible solutions. Also illustrated by the results are the introduction of chromatic aberration and the use of a slit to obtain desired end conditions. The results demonstrate that good agreement was obtained between the results predicted by the program and actual experimental results obtained with the 600 MeV cyclotron at the NASA Space Radiation Effects Laboratory.

Langley Research Center,
National Aeronautics and Space Administration,
Hampton, Va., November 12, 1971.

REFERENCES

1. Steffen, Klaus G.: High Energy Beam Optics. Interscience Publ., c.1965.

2. Good, Robert H.; and Piccioni, Oreste: Analog Computer for Charged Particle Trajectories. Rev. Sci. Instrum., vol. 31, no. 10, Oct. 1960, pp. 1035-1039.

3. Rampy, John M; and Berry, Donald T.: Determination of Stability Derivatives From Flight Test Data by Means of High Speed Repetitive Operation Analog Matching. FTC-TDR-64-8, U.S. Air Force, May 1964. (Available from DDC as AD 440 785.)

4. Colonias, John S.: TRACE: An On-Line Beam Transport Design Program. UCRL-18816 Preprint (AEC Contract No. W-7405-eng-48), Lawrence Radiat. Lab., Univ. of California, Apr. 1969.

5. Baker, W. F.: A Computer Program To Optimize Magnets in a Beam Transport System. Nucl. Instrum. Methods, vol. 20, Jan. 1963, pp. 55-57.

6. Hawkes, P. W.: Quadrupole Optics. Springer Tracts in Modern Physics, Vol. 42, G. Höhler, ed., Springer Verlag, 1966.

7. White, Ellis: Eastern Simulation Council Meeting. Simulation, vol. 12, no. 2, Feb. 1969, pp. 53-56.

8. Mitchell, Baker A., Jr.: A Hybrid Analog-Digital Parameter Optimizer for ASTRAC II. AFIPS Conference Proceedings, Vol. 25 – Spring Joint Computer Conference, Spartan Books, Inc., 1964, pp. 271-285.

9. Bard, Yonathan: Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems. SIAM J. Numerical Anal., vol. 7, no. 1, Mar. 1970, pp. 157-186.

10. Parrish, Russell V.: Parameter Identification Using a Creeping-Random-Search Algorithm. NASA TN D-6533, 1971.

11. Bowker, Albert H.; and Lieberman, Gerald J.: Engineering Statistics. Prentice-Hall, Inc., c.1959.

12. Steel, Robert G. D.; and Torrie, James H.: Principles and Procedures of Statistics. McGraw-Hill Book Co., Inc., 1960.

TABLE 1.- VARIATION OF NUMBER OF POSSIBLE DIRECTIONS WITH NUMBER OF PARAMETERS

Number of parameters	Possible directions (*)
1	2
2	8
3	26
4	80
5	242
6	728
7	2186

* $3^n - 1$ possible directions.

TABLE 2.- PARALLEL-TO-PARALLEL CASE

Parameter	Analog	Digital	
		No aberration	5% aberration
K_1, m^{-1}	0.9110	1.1525	1.1525
K_2, m^{-1}	-0.9398	-1.5665	-1.5665
K_3, m^{-1}	0.8812	0.9125	0.9125
K_4, m^{-1}	-0.6596	-0.8250	-0.8250
D_2, m	3.21	4.0125	4.0125
C	3.2747	2.2222	3.2117
$\overline{x'}(s_f)$	1.4965×10^{-3}	1.0562×10^{-3}	1.2758×10^{-3}
$\overline{z'}(s_f)$	1.59914×10^{-3}	1.0949×10^{-3}	1.1208×10^{-3}
l_1, m	0.7176		
l_2, m	0.7176		
l_3, m	0.6858		
l_4, m	0.6858		
D_1, m	0.1969		
D_3, m	0.2286		
N, percent	0		
W_{dif}	1.0		
W_N	0		

TABLE 3.- SLIT IN VERTICAL PLANE

[Parallel-to-parallel case; no chromatic aberration]

Parameter	Without slit	With slit
K_1, m^{-1}	1.165	1.165
K_2, m^{-1}	-1.554	-1.554
K_3, m^{-1}	0.938	0.938
K_4, m^{-1}	-0.837	-0.837
D_2, m	3.747	3.747
C	2.467	1.967
$\overline{x'}(s_f)$	1.119×10^{-3}	1.027×10^{-3}
$\overline{z'}(s_f)$	1.276×10^{-3}	6.566×10^{-4}
N, percent	0	63.33
l_1, m	0.7176	
l_2, m	0.7176	
l_3, m	0.6858	
l_4, m	0.6858	
D_1, m	0.1969	
D_3, m	0.2286	
W_{dif}	1.0	
r, m	0.3	
x_{sl}, m	0.045	

TABLE 4.- POINT-TO-PARALLEL CASE

[No chromatic aberration; no focal-point constraints]

Parameter	Value
K_1, m^{-1}	1.265625
K_2, m^{-1}	-1.065625
K_3, m^{-1}	1.118750
K_4, m^{-1}	-0.934375
D_2, m	4.28750
$\overline{x'}(s_f)$	1.1102×10^{-4}
$\overline{z'}(s_f)$	2.2458×10^{-4}
C	3.14786
l_1, m	0.8128
l_2, m	0.8128
l_3, m	0.5588
l_4, m	0.5588
D_1, m	0.1016
D_3, m	0.0508
W_{dif}	1.0
W_N	0
x(0), m	0.001
x'(0)	0.01
z(0), m	0.001
z'(0)	0.01

TABLE 5.- PARALLEL-TO-POINT CASE

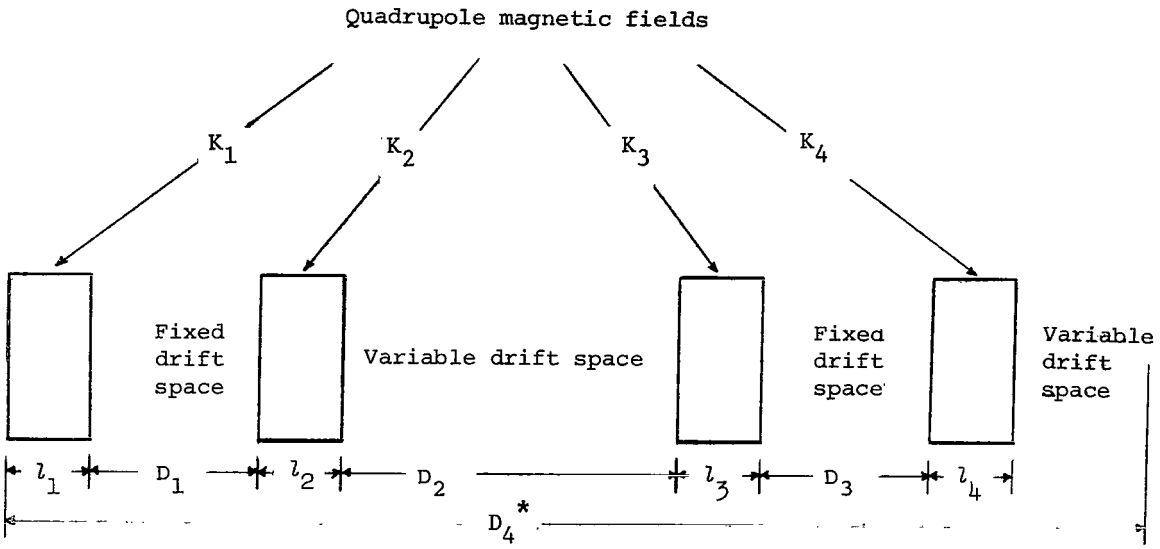
[No chromatic aberration; no focal-point constraints]

Parameter	Value
K_1, m^{-1}	-0.2340
K_2, m^{-1}	0.4180
K_3, m^{-1}	-1.0650
K_4, m^{-1}	1.4468
D_2, m	4.7875
D_4, m	9.0063
C	4.2084×10^{-7}
l_1, m	0.5588
l_2, m	0.5588
l_3, m	0.8128
l_4, m	0.8128
D_1, m	0.0508
D_3, m	0.1016
W_{dif}	1.0
W_N	0
x(0), m	0.055
x'(0)	0.000235
z(0), m	0.055
z'(0)	0.000235

TABLE 6.- POINT-TO-POINT CASE

[No chromatic aberration; no focal-point constraints]

Parameter	Value
K_1, m^{-1}	0.4000
K_2, m^{-1}	1.3000
K_3, m^{-1}	-1.2000
K_4, m^{-1}	1.2499
D_2, m	4.7000
D_4, m	11.0656
C	6.453×10^{-6}
l_1, m	0.7176
l_2, m	0.7176
l_3, m	0.6858
l_4, m	0.6858
D_1, m	0.1969
D_3, m	0.2286
W_{dif}	1.0
W_N	0
x(0), m	0.001
x'(0)	0.01
z(0), m	0.001
z'(0)	0.01



* D_4 is a variable in the parallel-to-point and point-to-point cases

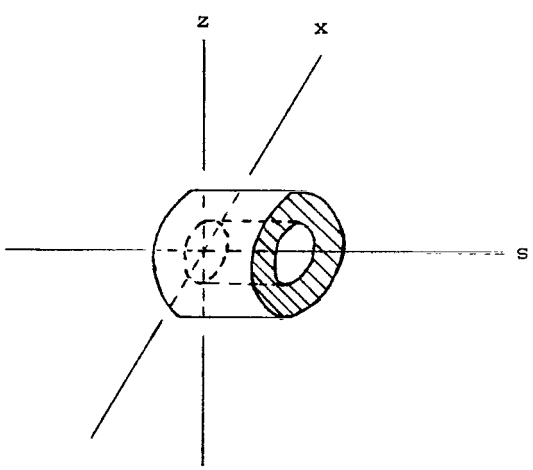
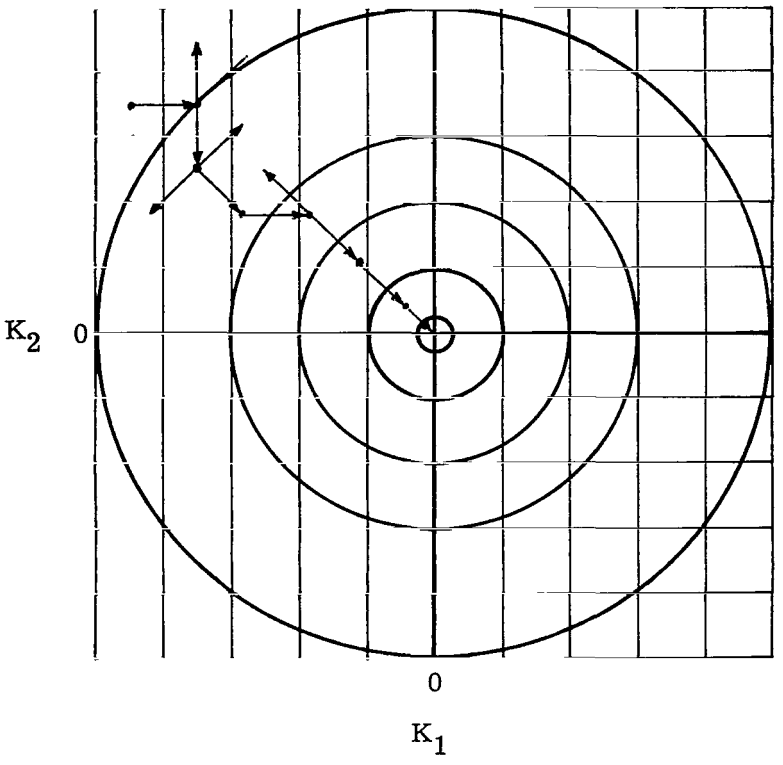
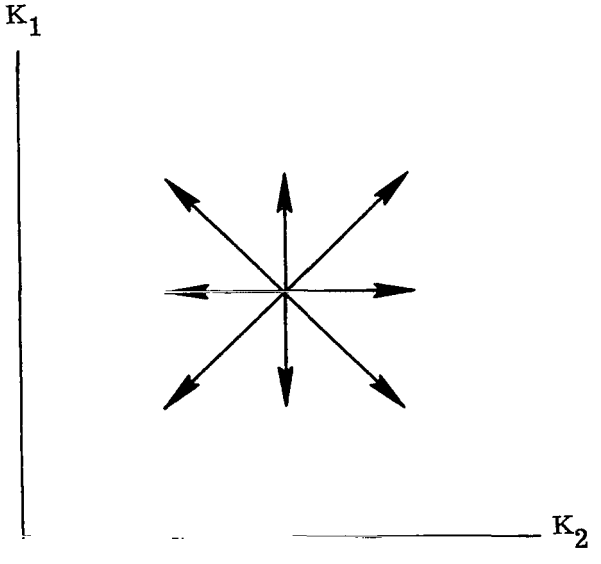


Figure 1.- Beam transport system.



(a) Two-parameter contour plot.



(b) Parameter directions.

Figure 2.- Illustration of basic algorithm.

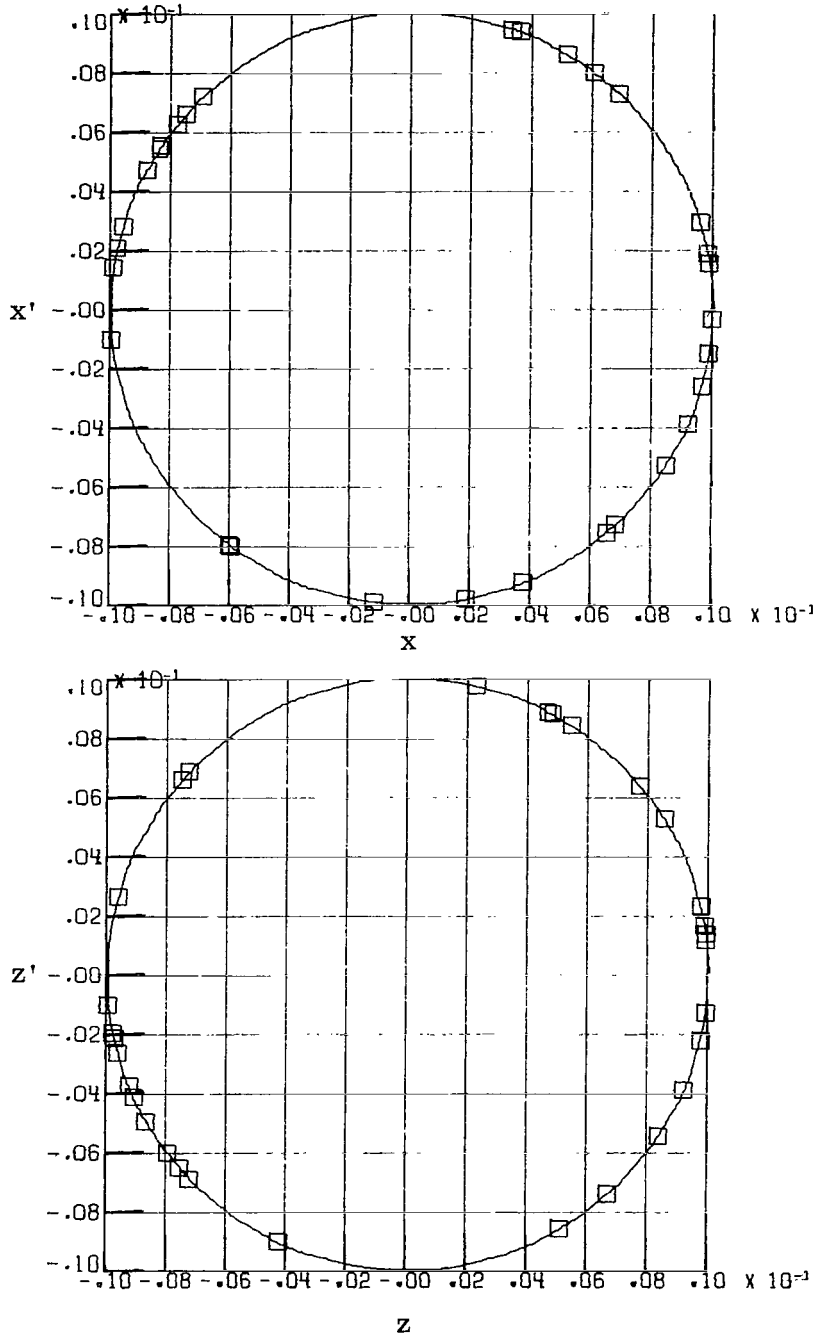


Figure 3.- Standard distribution for parallel initial conditions.

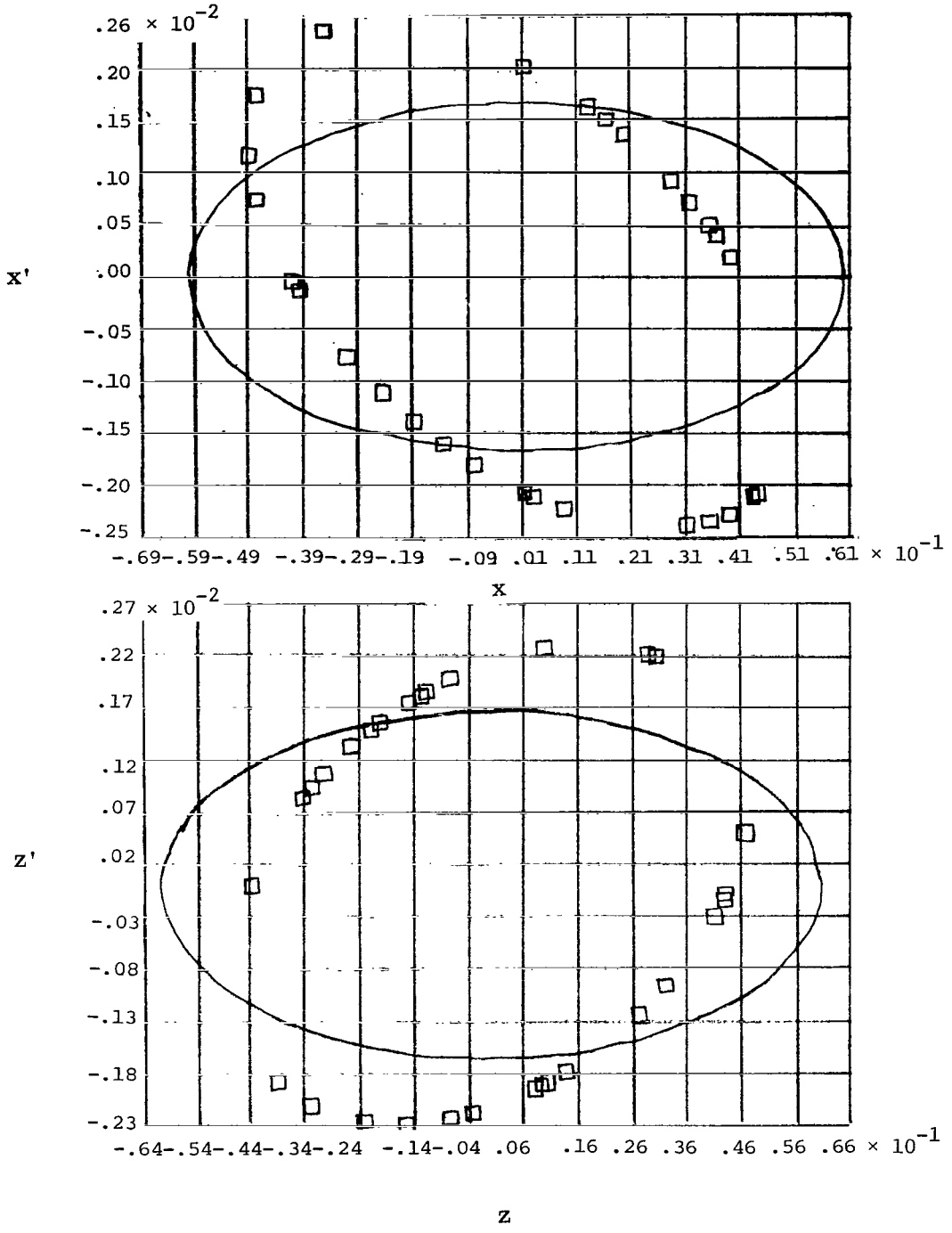
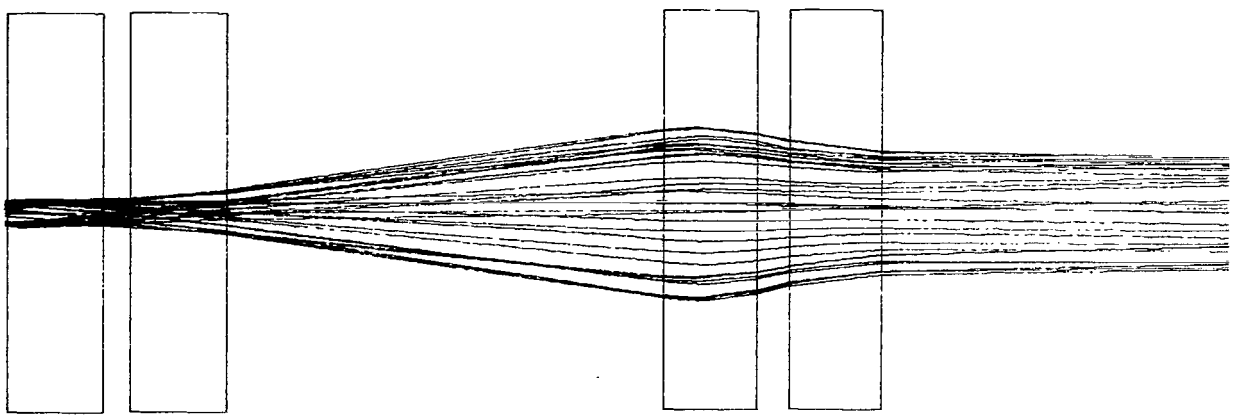
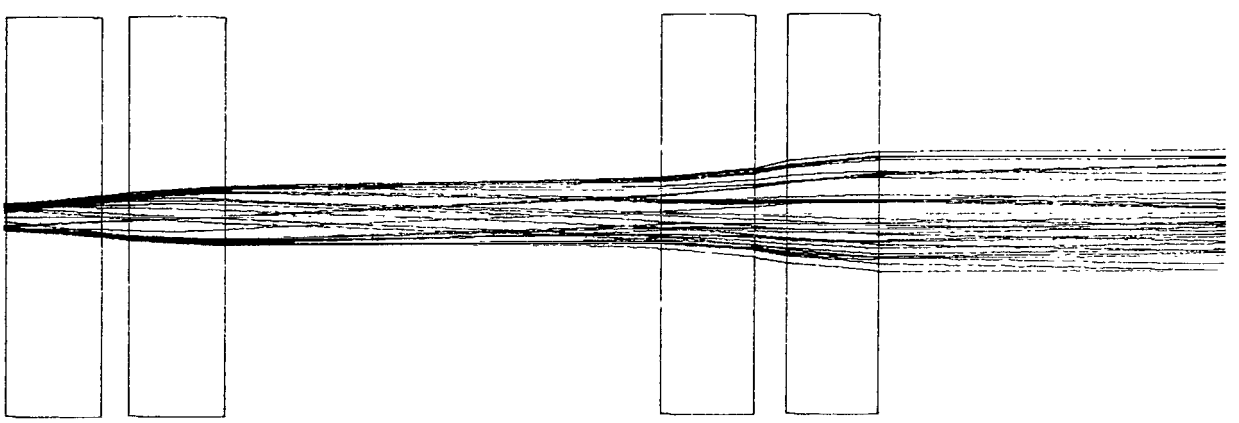


Figure 4.- End conditions for analog case. $K_1 = 0.9110\ m^{-1}$; $K_2 = -0.9398\ m^{-1}$; $K_3 = 0.8812\ m^{-1}$; $K_4 = -0.6596\ m^{-1}$; and $D_2 = 3.21$ m.



BOTTOM VIEW



SIDE VIEW

Figure 5.- Trace of analog case. $K_1 = 0.9110\ m^{-1}$; $K_2 = -0.9398\ m^{-1}$; $K_3 = 0.8812\ m^{-1}$; $K_4 = -0.6596\ m^{-1}$; and $D_2 = 3.21$ m.

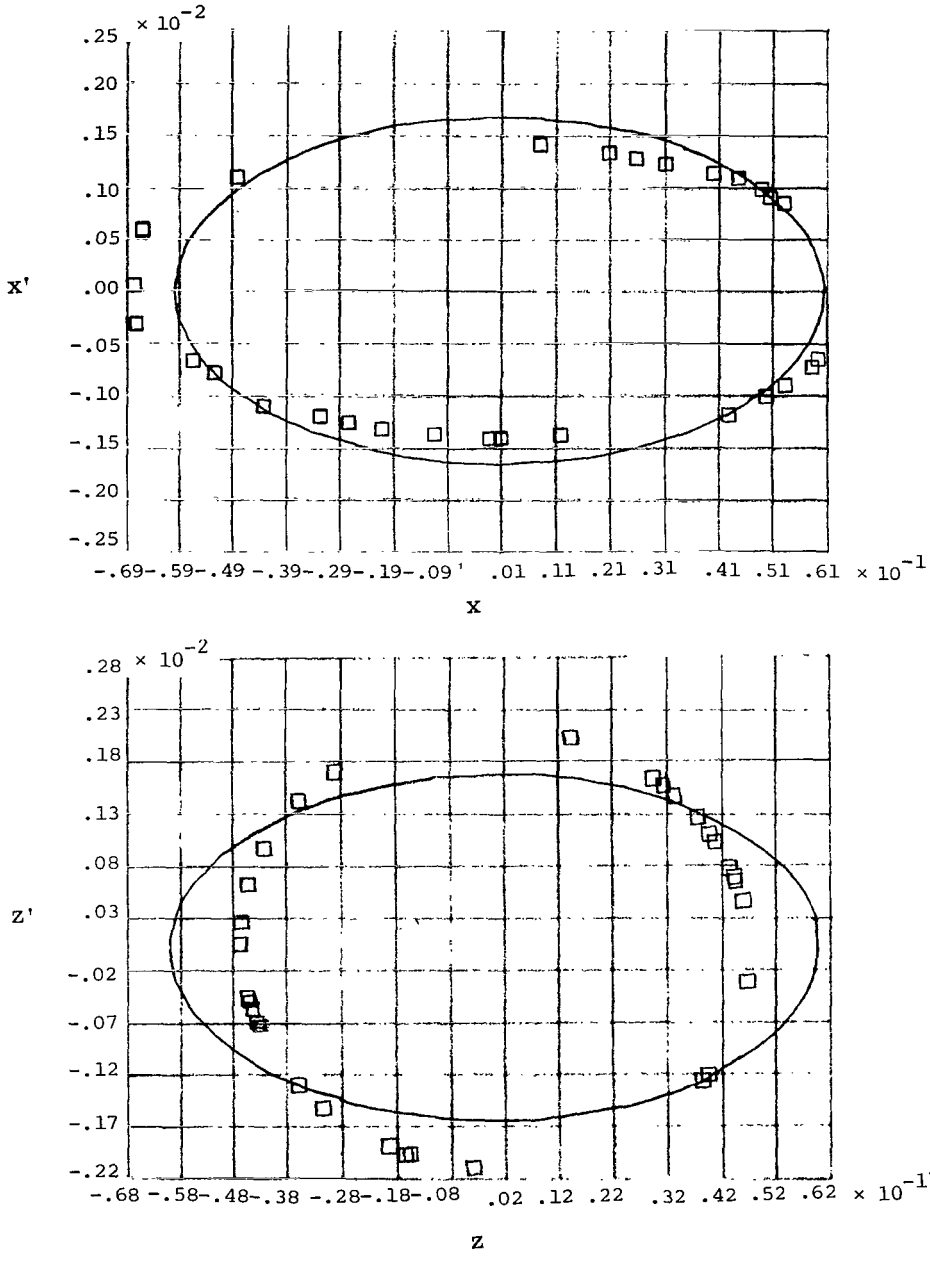
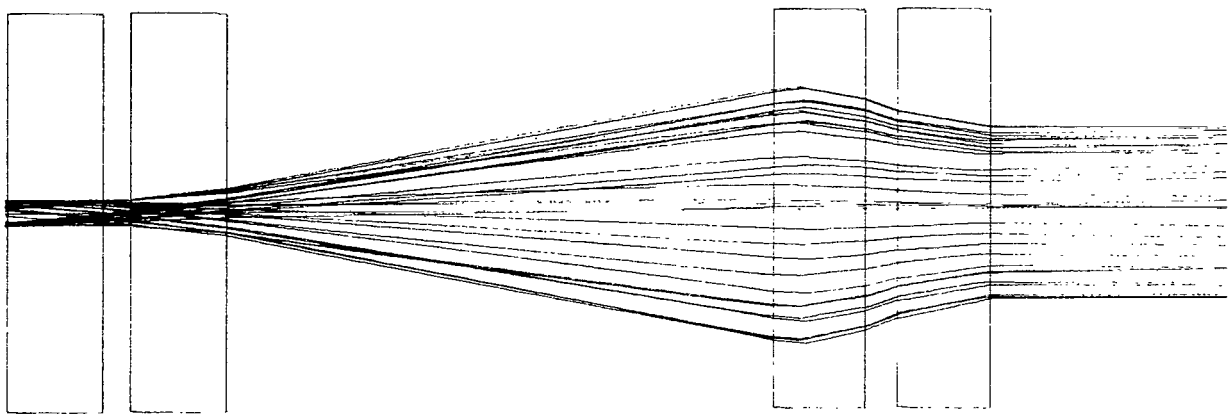


Figure 6.- End conditions for digital case. $K_1 = 1.1525\ m^{-1}$; $K_2 = -1.5665\ m^{-1}$; $K_3 = 0.9125\ m^{-1}$; $K_4 = -0.8250\ m^{-1}$; and $D_2 = 4.0125$ m.



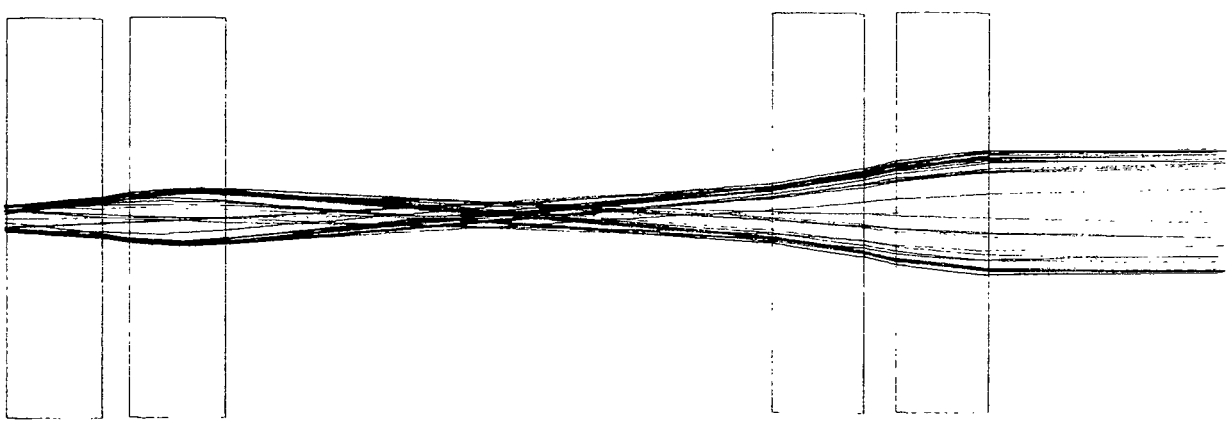


Figure 7.- Trace of digital case. $K_1 = 1.1525\ m^{-1}$; $K_2 = -1.5665\ m^{-1}$; $K_3 = 0.9125\ m^{-1}$; $K_4 = -0.8250\ m^{-1}$; and $D_2 = 4.0125$ m.

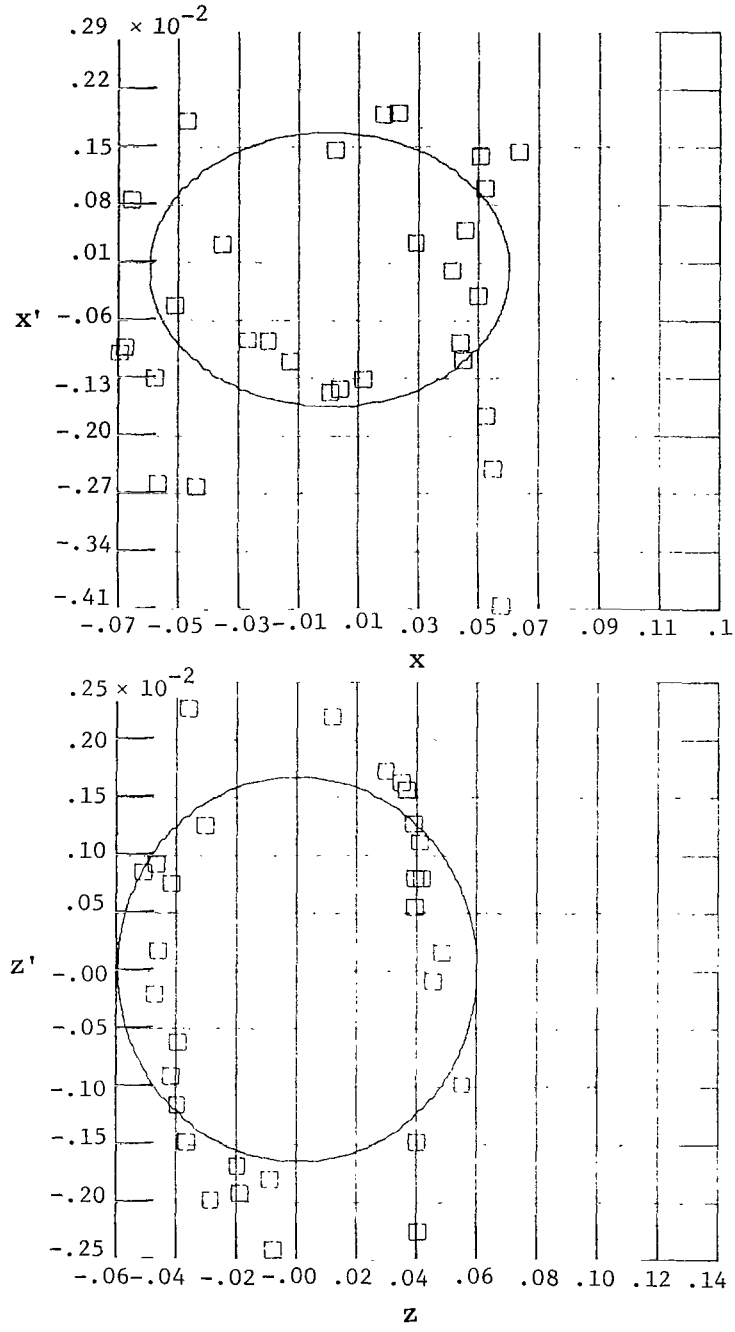
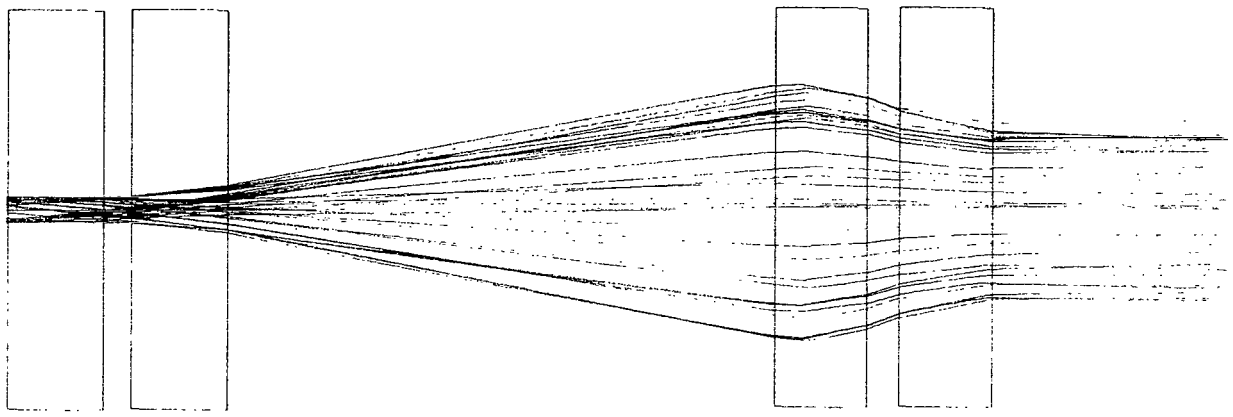
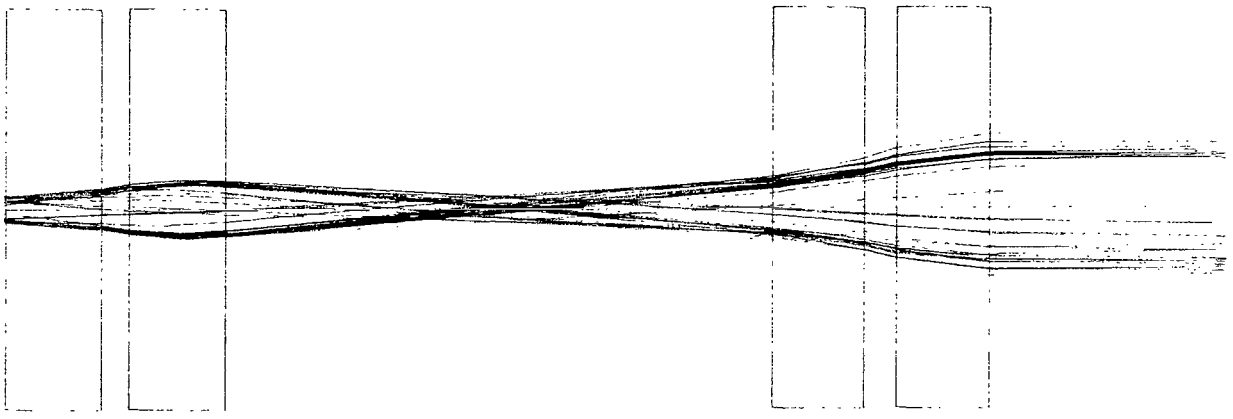


Figure 8.- End conditions for digital case with 5% chromatic aberration. $K_1 = 1.1525\ m^{-1}$; $K_2 = -1.5665\ m^{-1}$; $K_3 = 0.9125\ m^{-1}$; $K_4 = -0.8250\ m^{-1}$; and $D_2 = 4.0125$ m.



BOTTOM VIEW



SIDE VIEW

Figure 9.- Trace of digital case with 5% chromatic aberration. $K_1 = 1.1525\ m^{-1}$; $K_2 = -1.5665\ m^{-1}$; $K_3 = 0.9125\ m^{-1}$; $K_4 = -0.8250\ m^{-1}$; and $D_2 = 4.0125$ m.

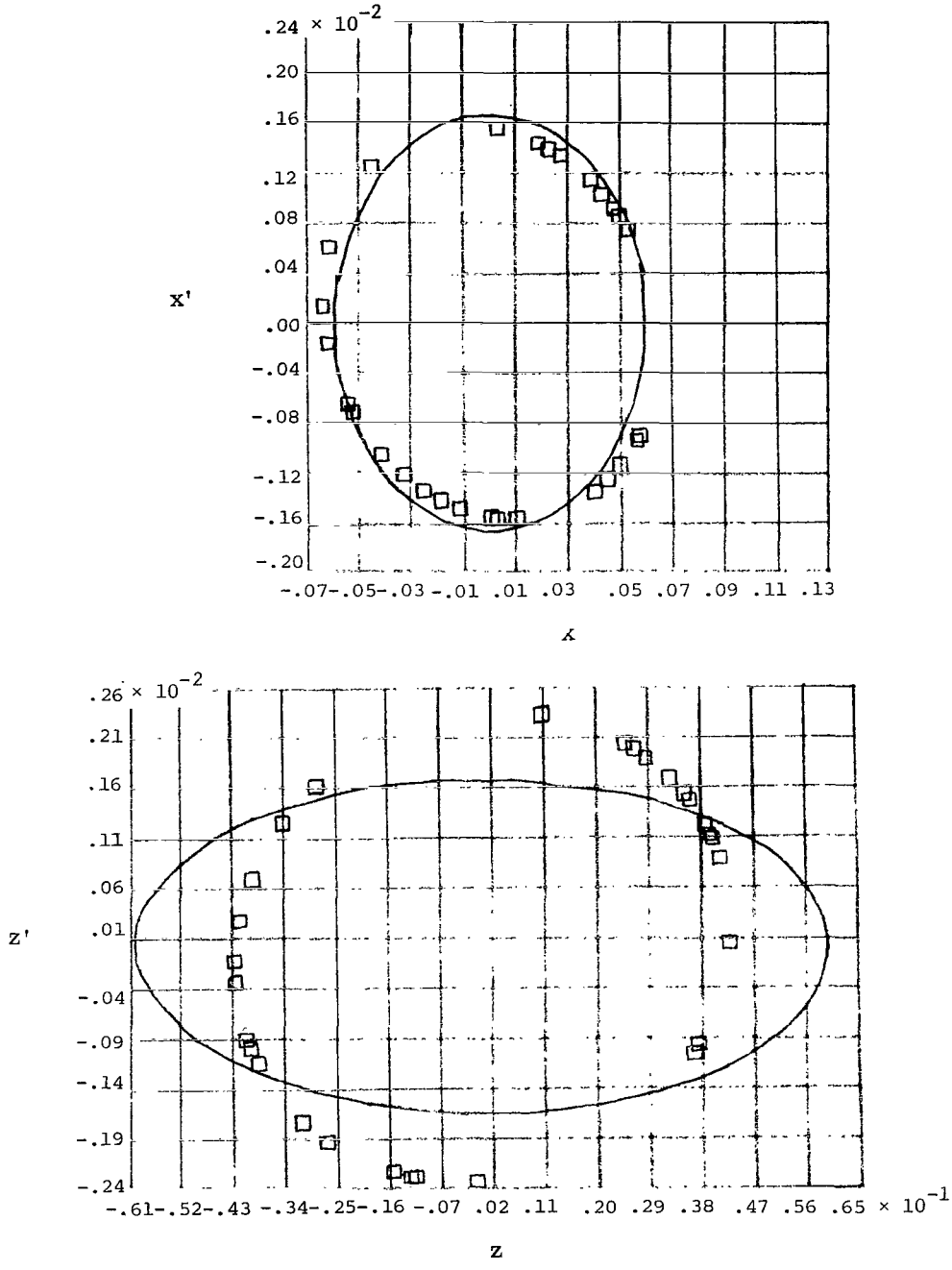
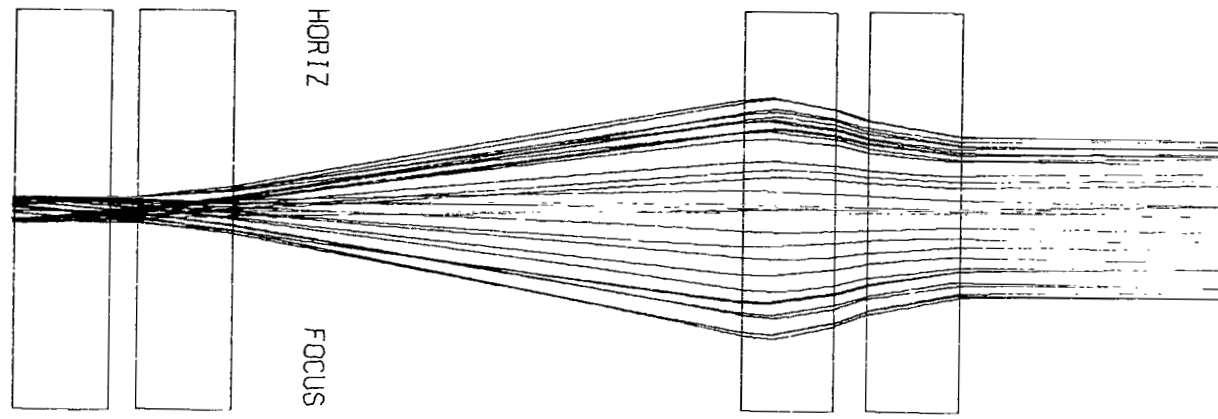
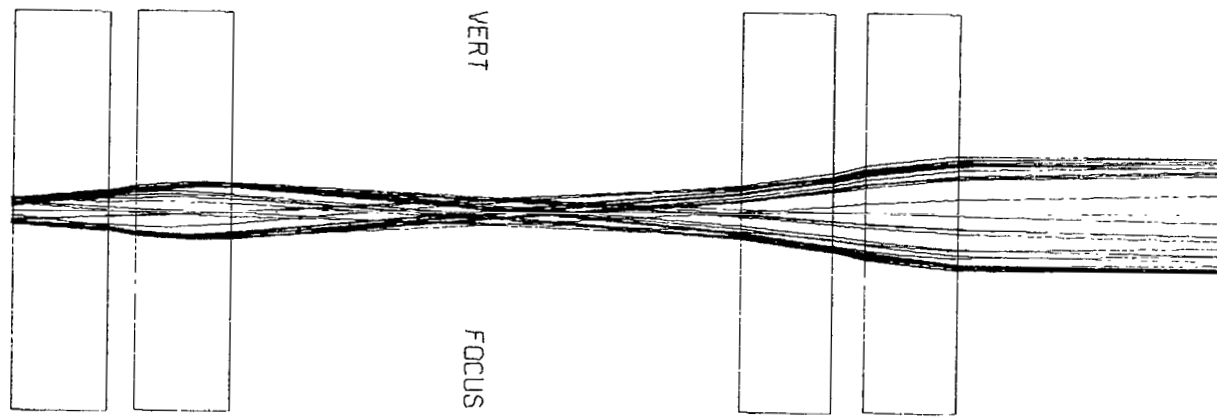


Figure 10.- End conditions for system prior to introduction of slits. $K_1 = 1.165\ m^{-1}$; $K_2 = -1.554\ m^{-1}$; $K_3 = 0.938\ m^{-1}$; $K_4 = -0.837\ m^{-1}$; and $D_2 = 3.747$ m.



BOTTOM VIEW



SIDE VIEW

Figure 11.- Trace of system prior to introduction of slits.
$K_1 = 1.165\ m^{-1}$; $K_2 = -1.554\ m^{-1}$; $K_3 = 0.938\ m^{-1}$;
$K_4 = -0.837\ m^{-1}$; and $D_2 = 3.747$ m.

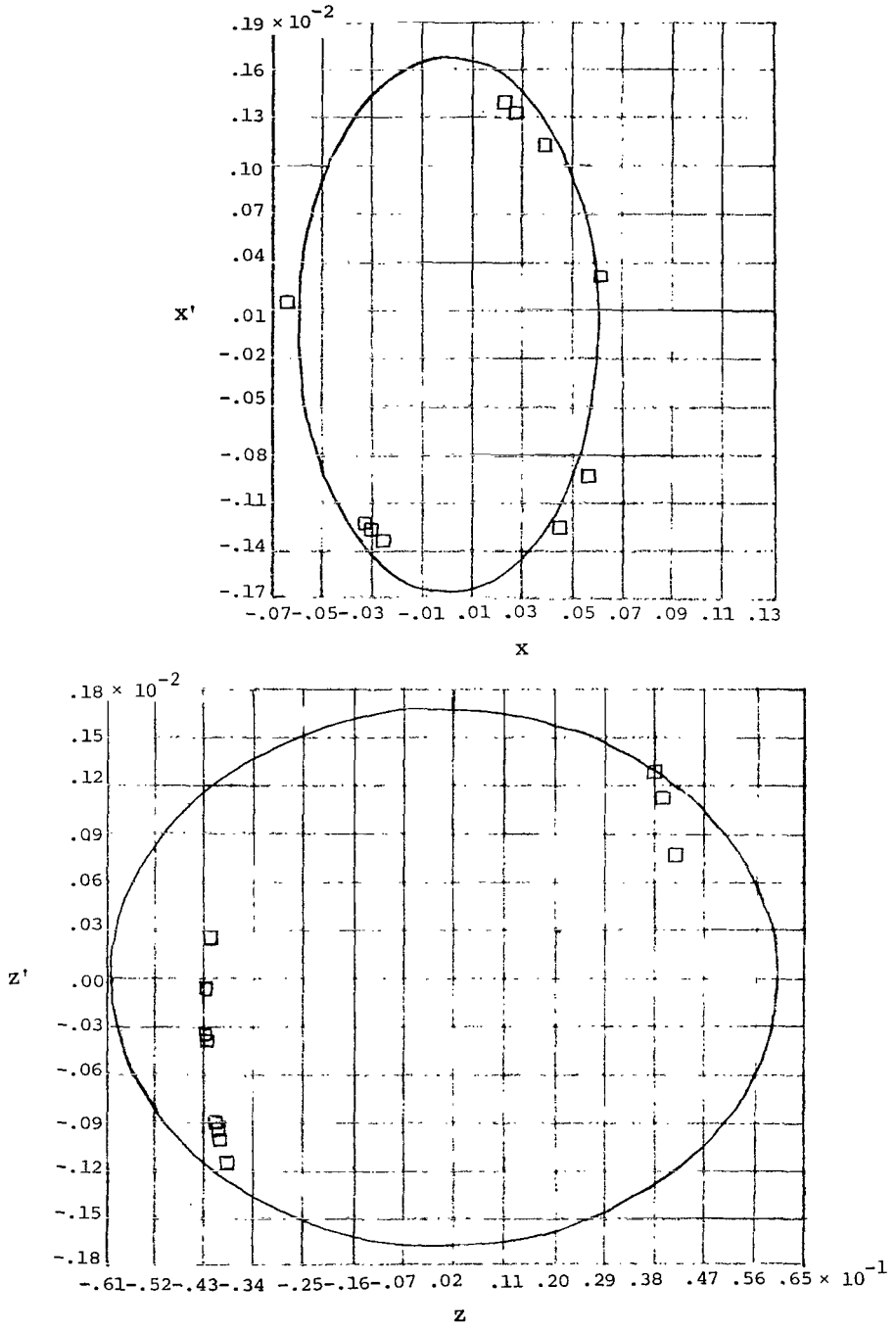
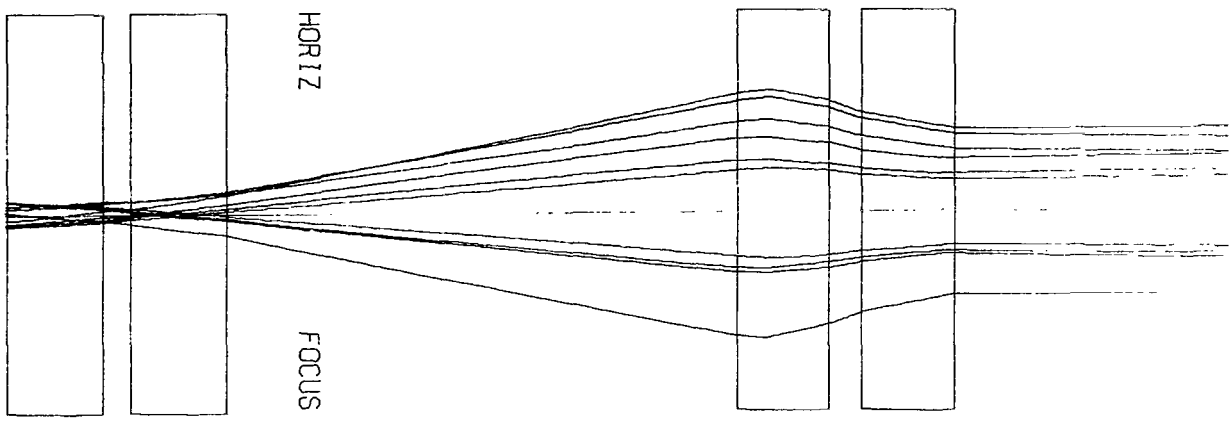
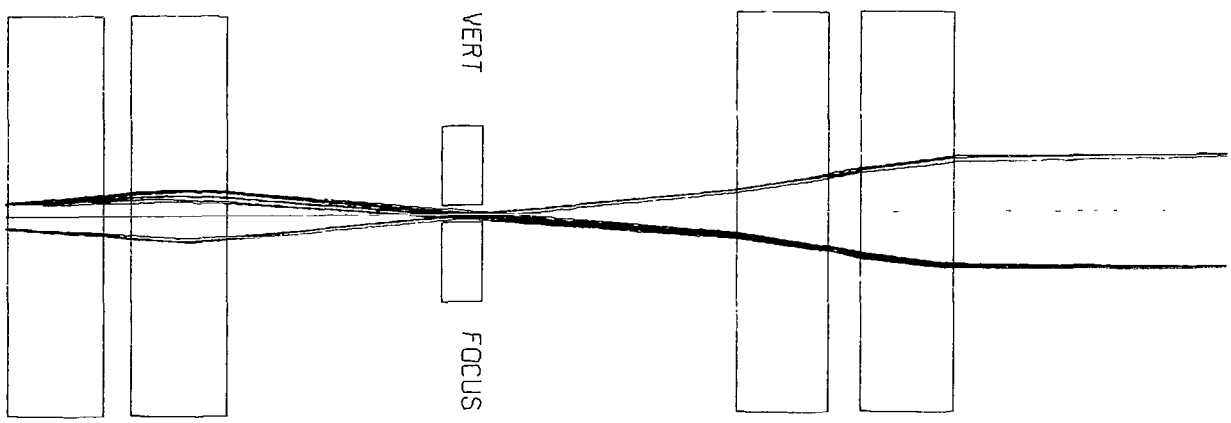


Figure 12.- End conditions for system with slits. $K_1 = 1.165\ m^{-1}$; $K_2 = -1.554\ m^{-1}$; $K_3 = 0.938\ m^{-1}$; $K_4 = -0.837\ m^{-1}$; and $D_2 = 3.747$ m.

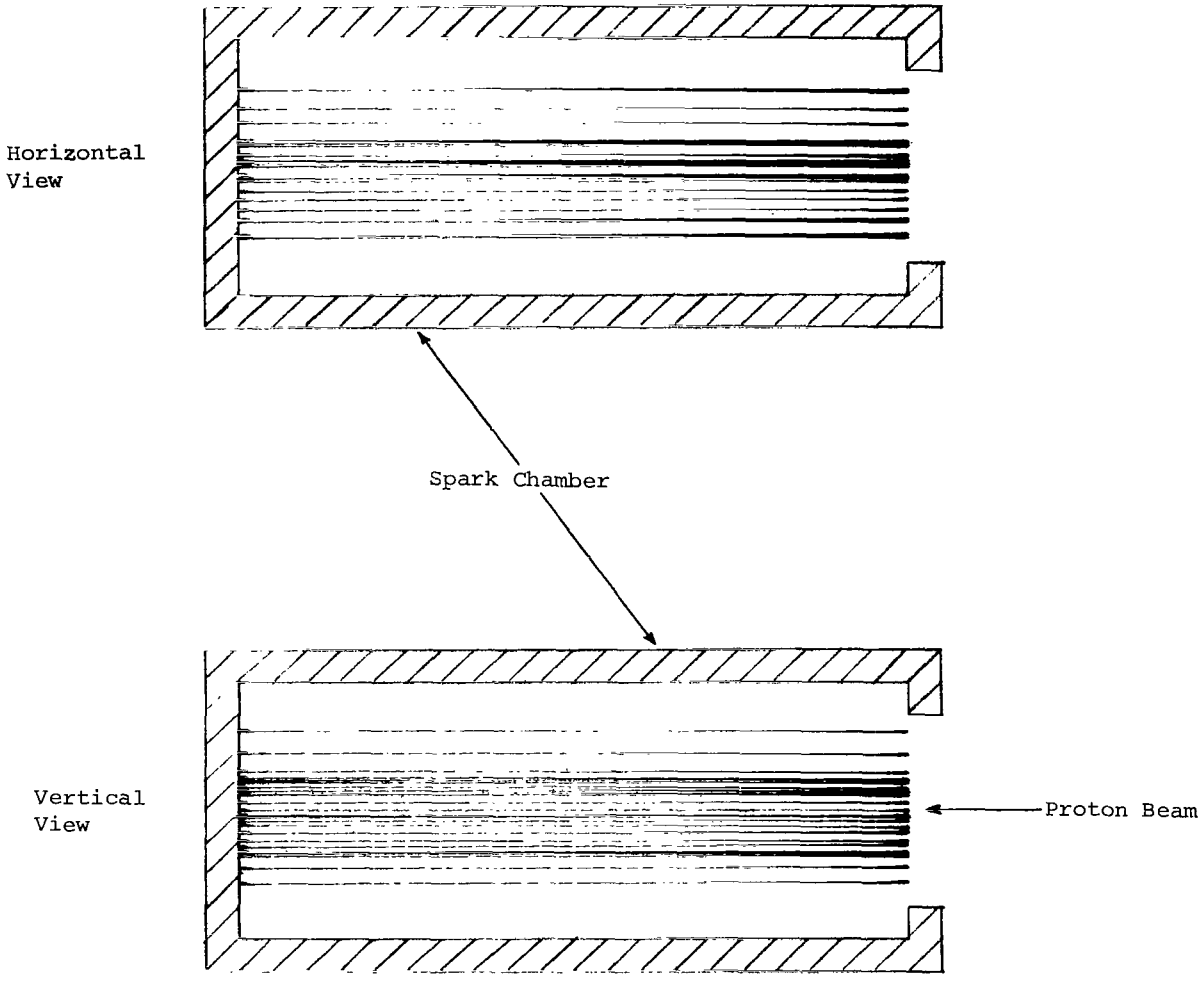


BOTTOM VIEW



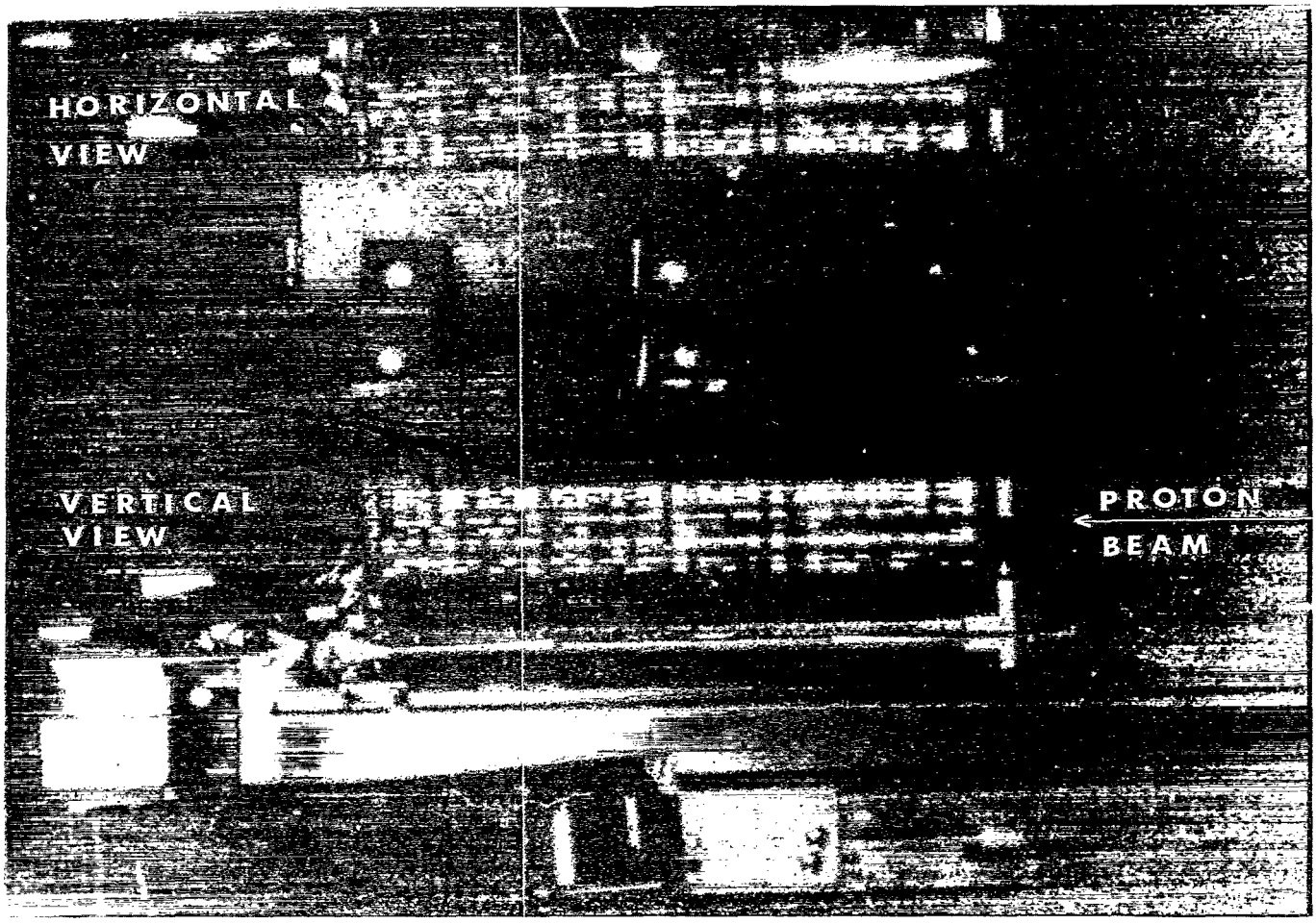
SIDE VIEW

Figure 13.- Trace of system with slits. $K_1 = 1.165\ m^{-1}$; $K_2 = -1.554\ m^{-1}$; $K_3 = 0.938\ m^{-1}$; $K_4 = -0.837\ m^{-1}$; and $D_2 = 3.747$ m.



(a) Computer-generated curves.

Figure 14.- Comparison of computer results and actual experiment.



L-71-7119

(b) Results of experiment.

Figure 14.- Concluded.

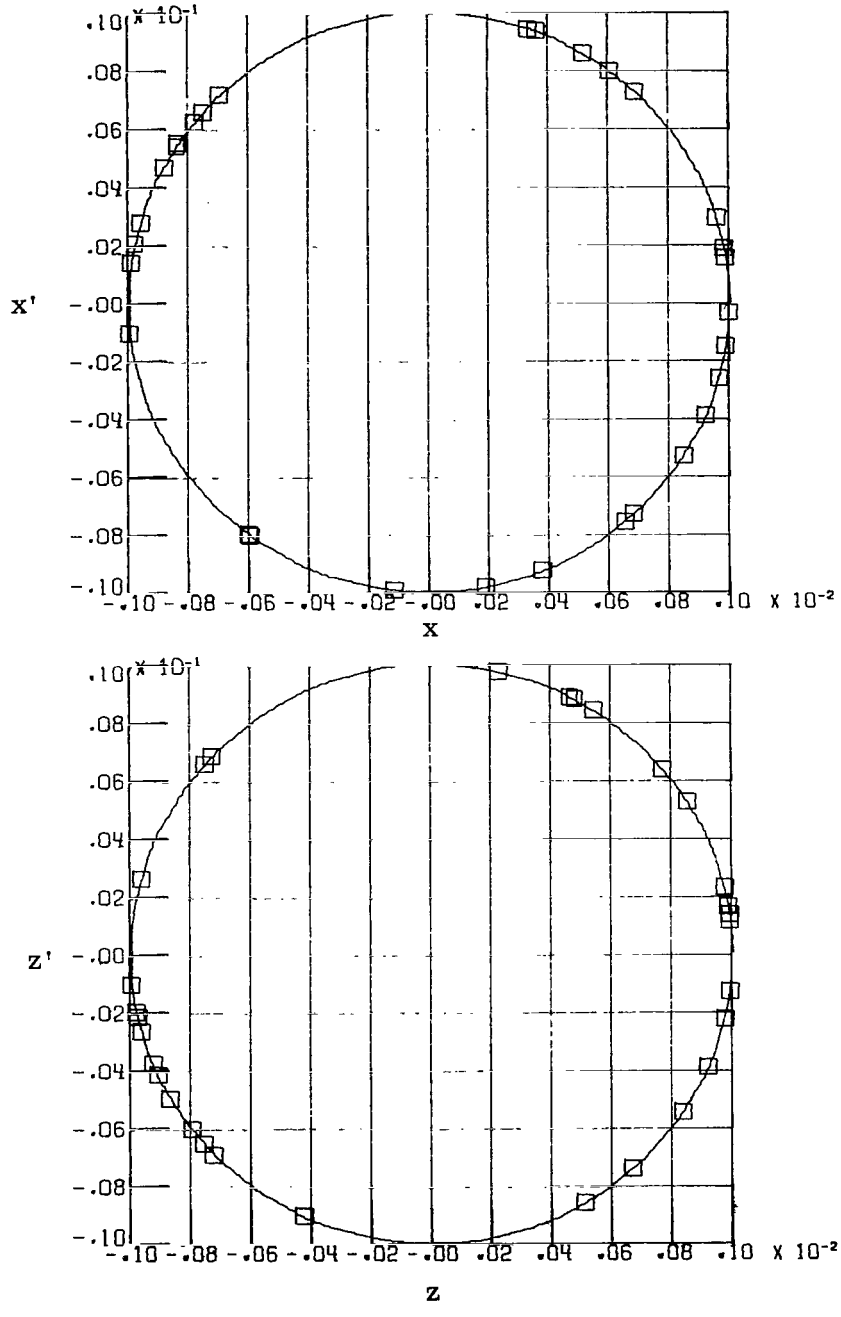


Figure 15.- Distribution for point-to-parallel case initial conditions.

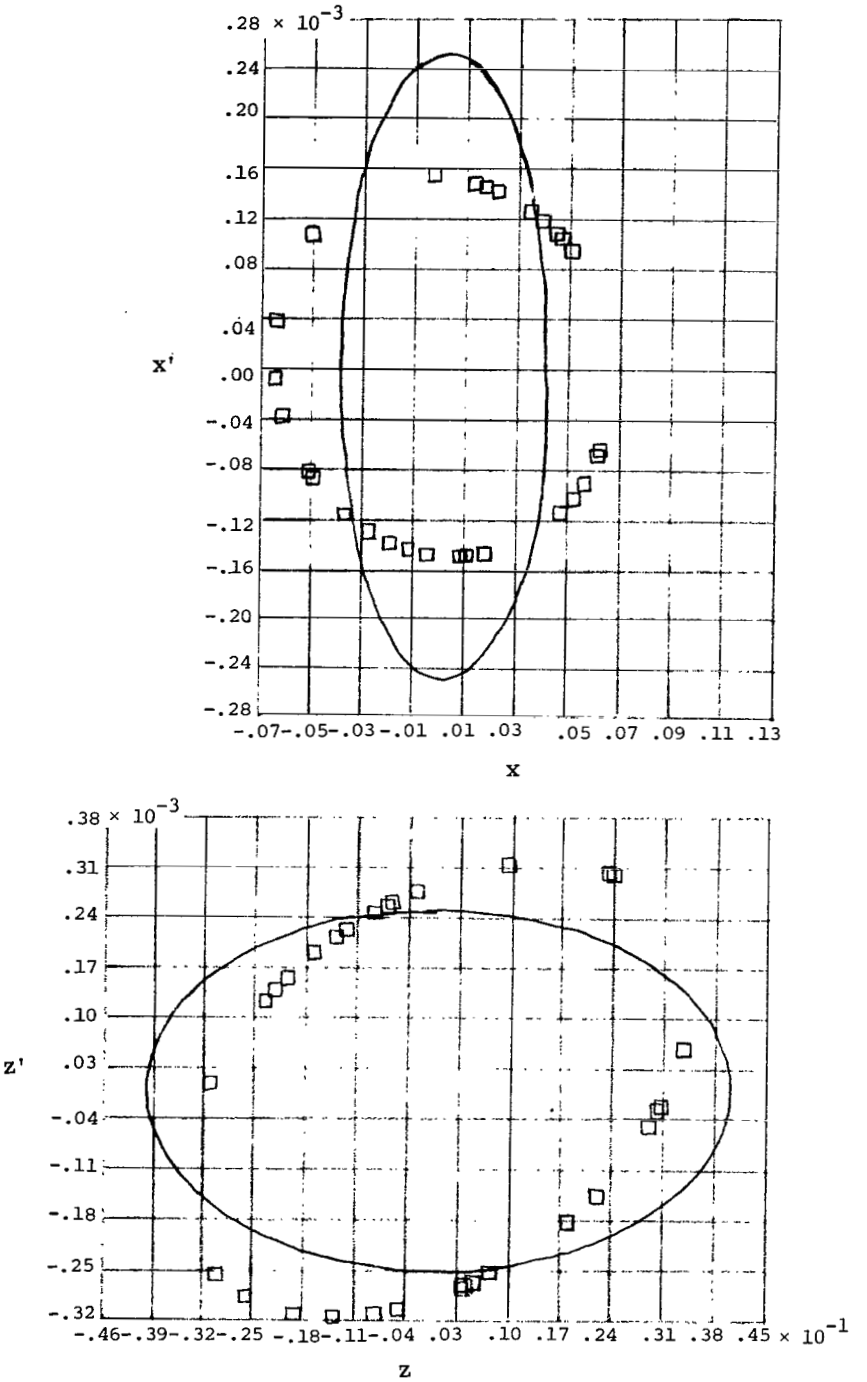
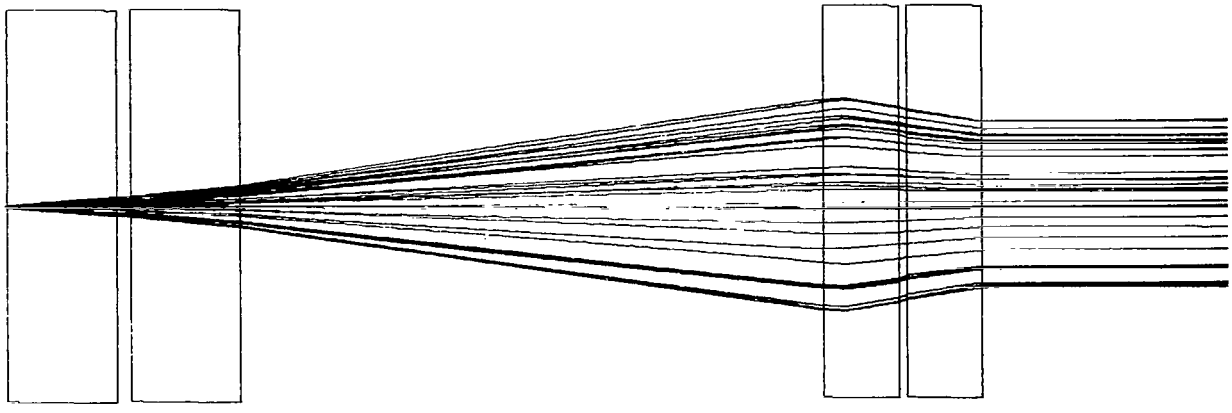


Figure 16.- End conditions for point-to-parallel case.



BOTTOM VIEW

SIDE VIEW

Figure 17.- Trace of point-to-parallel case. $K_1 = 1.265625\ m^{-1}$; $K_2 = -1.065625\ m^{-1}$; $K_3 = 1.118750\ m^{-1}$; $K_4 = -0.934375\ m^{-1}$; and $D_2 = 4.28750$ m.

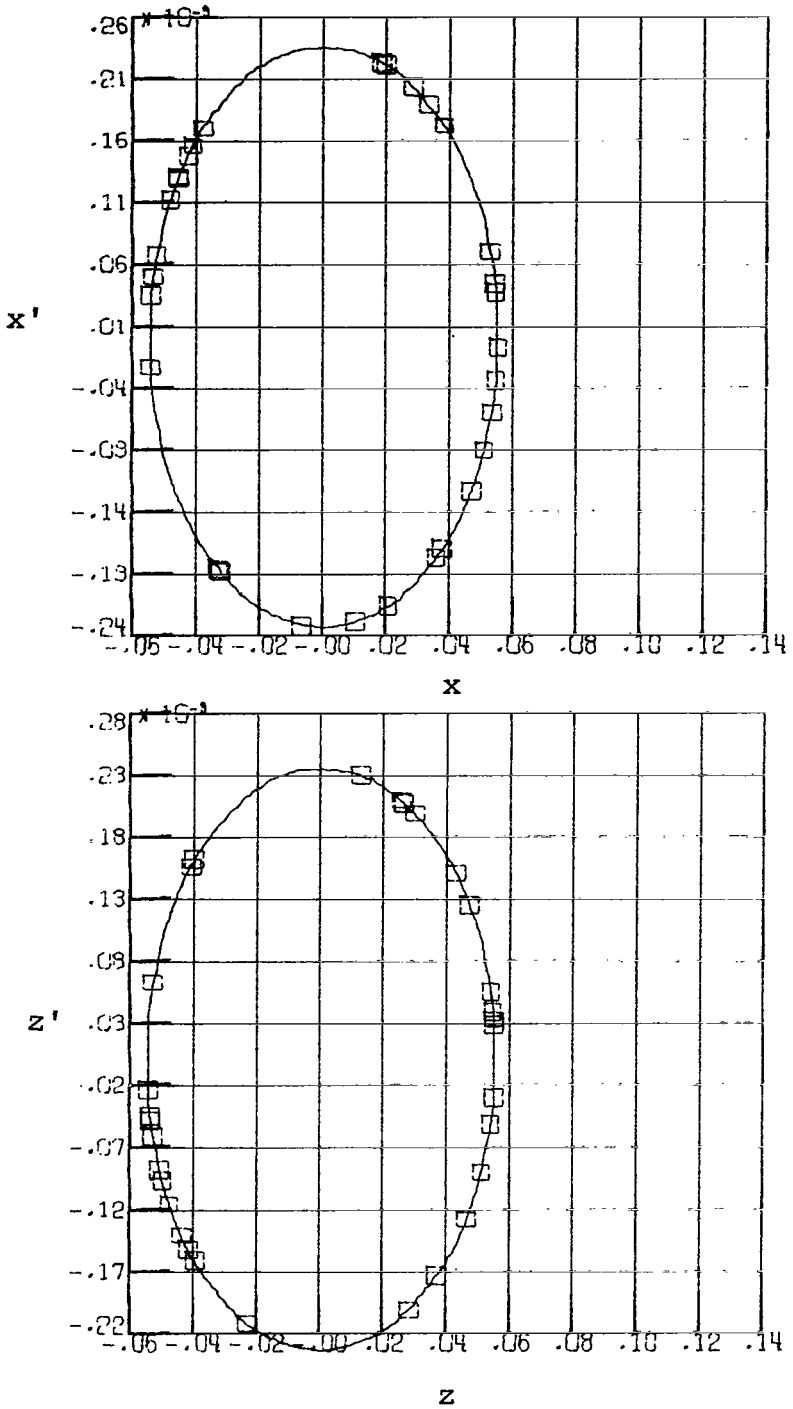
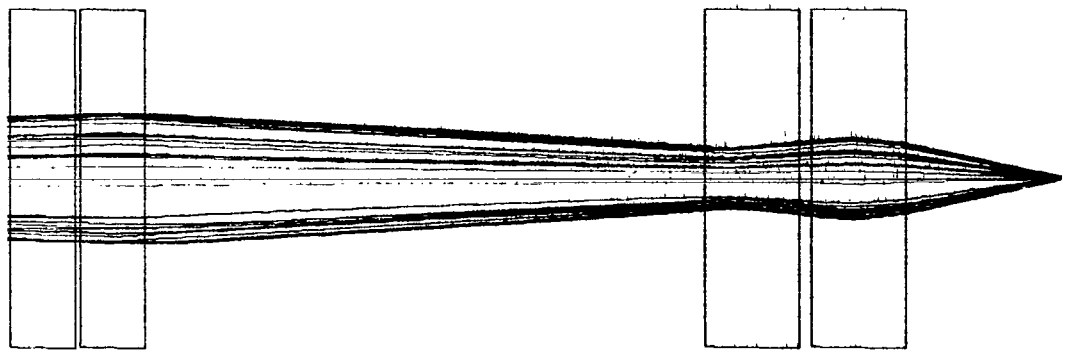
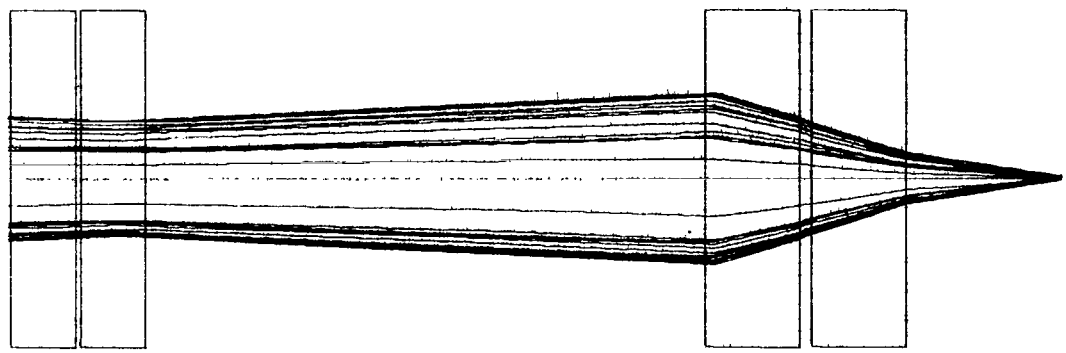


Figure 18.- Distribution for parallel-to-point case initial conditions.

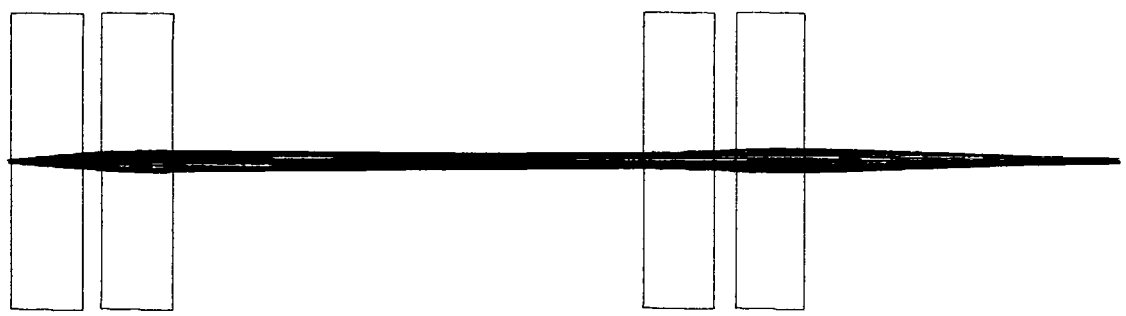


BOTTOM VIEW

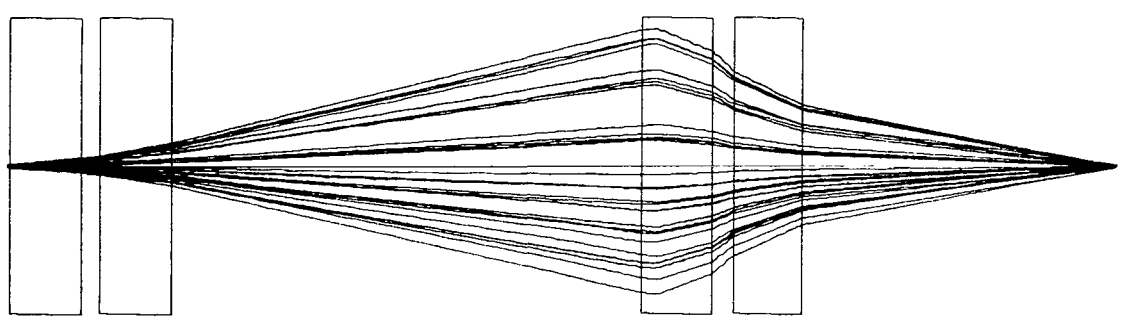


SIDE VIEW

Figure 19.- Trace of parallel-to-point case. $K_1 = -0.2340\ m^{-1}$; $K_2 = 0.4180\ m^{-1}$; $K_3 = -1.0650\ m^{-1}$; $K_4 = 1.4468\ m^{-1}$; $D_2 = 4.7875$ m; and $D_4 = 9.0063$ m.



BOTTOM VIEW



SIDE VIEW

Figure 20.- Trace of point-to-point case. $K_1 = 0.4000\ m^{-1}$; $K_2 = 1.3000\ m^{-1}$; $K_3 = -1.2000\ m^{-1}$; $K_4 = 1.2499\ m^{-1}$; $D_2 = 4.7000$ m; and $D_4 = 11.0656$ m.

NASA-Langley, 1971 — 19 L-7930

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546